



```

DR N-PSDB_N80603.
PT Restriction fragment length polymorphism genotyping test -
PT for chickens, using erythrocyte DNA fragments and hybridisation
PT probe derived from histocompatibility complex antigen.
PS Disclosure: p. French.
CC Probe F10 is used in RFLP analysis of chicken erythrocyte-derived
CC DNA. It corresponds to a MHC class I antigen and was obtained from
CC the total mRNA of different tissues of chicken.
SQ See also N80602 and N82429.
Sequence 345 AA.

Query Match
Best Local Similarity 41.0%; Score 97; DB 1; Length 345;
Matches 16; Conservative 8; Mismatches 12; Indels 3; Gaps 3

Db 161 ptkirweeseperpk-nyleetcwvllry-veygxae 197
164 PDAANYNDINDVKRLKPGYLEAT-VDMFRRYKVPDGPKE 201

RESULT 4
ID R42426 standard; Protein; 350 AA.
AC R42426.
DT 24-MAY-1994 (first entry)
DE Bovine rod transducin.
KW Taste modifying agent; ligand; antiligand; binding activity; taste;
KW taste receptor cells; sweet; bitter; salty; sour; gustducin.
OS Bos taurus.
PN M09321337-A.
PD 28-OCT-1993.
PF 08-APR-1993; U03279.
PR 09-APR-1992; US-8668353.
PA (MARG/) MARGOLSKEE R F.
PI Margolskee RF:
DR WPI: 93-351746/44.
PT New gustducin alpha subunit protein - used for identifying taste
PT modifying agents which mimic or inhibit sweet, bitter, salty or
PT sour tastes
PS Disclosure: Page 34; 50pp; English.
PS The sequence is that of bovine rod transducin. The sequence was
CC compared to that of bovine cone transducin and a rat novel taste
CC cell specific guanine nucleotide binding protein, gustducin, alpha
CC subunit. Fragments of the protein possessing at least one ligand/
CC antiligand binding activity or immunological property specific to
CC gustducin are taste modifying agents which can be delivered to
CC taste receptor cells to modify taste, e.g. mimic or inhibit sweet,
CC bitter, salty or sour tastes.
CC See also R42405-36.
SQ Sequence 350 AA.

Query Match
Best Local Similarity 4.5%; Score 95; DB 8; Length 350;
Matches 20; Conservative 17; Mismatches 23; Indels 5; Gaps 5;

Db 146 dsaaay-1sderlvtpgryptegdvlsrvsktt-gietqsfk-dlnfrmdivgqrs 202
165 DDAANYNDINDVKRL-KPGYLEATVDMFR-RYKVPDGPENEFNAFMDKDFAIIDIKS 222

QY 165 DDAANYNDINDVKRL-KPGYLEATVDMFR-RYKVPDGPENEFNAFMDKDFAIIDIKS 222
Db 203 erkvw 207
QY 223 THDMW 227

RESULT 5
ID R42424 standard; Protein; 354 AA.
AC R42424.
DT 24-MAY-1994 (first entry)
DE Rat gustducin alpha subunit.
KW taste modifying agent; ligand; antiligand; binding activity; taste;
KW taste receptor cells; sweet; bitter; salty; sour.
OS Rattus norvegicus.
PN M09321337-A.
PD 28-OCT-1993.
PF 08-APR-1993; U03279.
PR 09-APR-1992; US-8668353.
PA (MARG/) MARGOLSKEE R F.
PI Margolskee RF:
DR WPI: 93-351746/44.
PT New gustducin alpha subunit protein - used for identifying taste
PT modifying agents which mimic or inhibit sweet, bitter, salty or
PT sour tastes
PS Disclosure: Page 34; 50pp; English.
PS The sequence is that of bovine rod transducin. The sequence was
CC compared to that of bovine cone transducin and a rat novel taste
CC cell specific guanine nucleotide binding protein, gustducin, alpha
CC subunit. Fragments of the protein possessing at least one ligand/
CC antiligand binding activity or immunological property specific to
CC gustducin are taste modifying agents which can be delivered to
CC taste receptor cells to modify taste, e.g. mimic or inhibit sweet,
CC bitter, salty or sour tastes.
CC See also R42405-36.
SQ Sequence 350 AA.

```

Query Match	4.3%	Score 91	DB 3	Length 366
Best Local Similarity	50.0%	Pred. No.	2.69e+01	
Matches	11	Conservative	7	Mismatches 2
				Indels 2
				Gaps 2
Db	180	rraylegtcwewlrry-lengk	200	
		:-:	:-:	
Oy	179	KGYLEAT-VDMEFRXKVPDGR	199	
RESULT	7			
ID	R60101	standard	Protein	713 AA
AC	R60101			
DT	15-MAR-1995	(first entry)		
DE	Canine zona pelliculida C2P2			
KW	Canine: dog; zona pelliculida; ZP; C2P2; contraceptive: vaccine;			

```

Query Match          4.3%  Score 91;  DB 10;  Length 715;
Best Local Similarity 30.2%  Pred. No. 2.69e+01;
Matches 16;  Conservative 12;  Mismatches 23;  Indels 2;  Gaps 2.

Db 456 fmktykc-hysarddlintvqslppvavsvrpglalilqtypdksyilrpyg 507
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 44 FHMVEYVFRMSNAKMEIATK-DPINPIKODKKCKLRYVALTFPKKGIYMWYG 95

RESULT 9
ID P80911 standard; protein; 274 AA.
AC P80911.
DT 18-SEP-1990 (first entry)

```

[illegible]





\*\*\*\*\*  
MUSEUM  
(TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Msrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Thu Sep 17 16:57:21 1998; Maspar time 3.93 Seconds  
Tabular output not generated. 518.696 Million cell updates/sec

Title: >US-08-741-437-1  
Description: (1-289) from US08741437.rep  
Perfect Score: 2134  
Sequence: 1 MSGFSTEERAAPELSLEYRVF.....CESACTVPTDVKMFHOKN 289

Scoring table: PAM 150  
Gap 11

Searched: 77021 seqs, 7058996 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-issued  
1:5\_COMB 2:PCT9\_COMB 3:backfillset1

Statistics: Mean 31.457; Variance 133.821; scale 0.235

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	1035	48.5	286	2	PCT-US95-1	Sequence 3, Applicatio	2.04e-89
2	185	8.7	176	2	PCT-US95-1	Sequence 10, Applicati	2.28e-07
3	165	7.7	179	2	PCT-US95-1	Sequence 7, Applicatio	1.23e-05
4	154	7.2	164	2	PCT-US95-1	Sequence 6, Applicatio	1.07e-04
5	111	7.1	263	2	PCT-US95-1	Sequence 4, Applicatio	1.91e-02
6	118	5.5	175	2	PCT-US95-1	Sequence 5, Applicatio	1.45e+00
7	103	4.8	182	1	US-08-127-	Sequence 146, Applicat	2.06e+00
8	101	4.7	182	1	US-08-127-	Sequence 14, Applicati	5.83e+00
9	95	4.5	350	1	US-07-868-	Sequence 15, Applicati	4.14e+00
10	97	4.5	354	1	US-08-127-	Sequence 142, Applicat	9.73e+00
11	92	4.3	182	1	US-08-127-	Sequence 162, Applicat	9.73e+00
12	92	4.3	182	1	US-08-127-	Sequence 161, Applicat	9.73e+00
13	92	4.3	182	1	US-08-127-	Sequence 160, Applicat	9.73e+00
14	92	4.3	182	1	US-08-127-	Sequence 140, Applicat	9.73e+00
15	92	4.3	182	1	US-08-127-	Sequence 147, Applicat	9.73e+00
16	92	4.3	182	1	US-08-127-	Sequence 159, Applicat	9.73e+00
17	92	4.3	182	1	US-08-127-	Sequence 157, Applicat	9.73e+00
18	92	4.3	182	1	US-08-127-	Sequence 144, Applicat	9.73e+00
19	92	4.3	182	1	US-08-127-	Sequence 156, Applicat	9.73e+00
20	92	4.3	182	1	US-08-127-	Sequence 173, Applicat	9.73e+00
21	92	4.3	182	1	US-08-127-	Sequence 172, Applicat	9.73e+00
22	92	4.3	182	1	US-08-127-	Sequence 143, Applicat	9.73e+00
23	92	4.3	182	1	US-08-127-	Sequence 143, Applicat	9.73e+00

24	92	4.3	182	1	US-08-127-	Sequence 137, Applicat	9.73e+00
25	90	4.2	182	1	US-08-127-	Sequence 156, Applicat	1.36e+01
26	90	4.2	182	1	US-08-127-	Sequence 158, Applicat	1.36e+01
27	90	4.2	182	1	US-08-127-	Sequence 135, Applicat	1.36e+01
28	90	4.2	182	1	US-08-127-	Sequence 154, Applicat	1.36e+01
29	89	4.2	182	1	US-08-127-	Sequence 153, Applicat	1.61e+01
30	89	4.2	182	1	US-08-127-	Sequence 145, Applicat	1.61e+01
31	90	4.2	354	1	US-08-222-	Sequence 12, Applicati	1.36e+01
32	90	4.2	375	2	PCT-US93-0	Sequence 8, Applicatio	1.36e+01
33	90	4.2	1079	1	US-08-484-	Sequence 8, Applicatio	1.36e+01
34	90	4.2	1085	1	US-08-485-	Sequence 5, Applicatio	1.61e+01
35	90	4.2	1085	1	US-08-485-	Sequence 5, Applicatio	1.61e+01
36	89	4.2	1085	1	US-08-485-	Sequence 5, Applicatio	1.61e+01
37	89	4.2	1085	1	US-08-485-	Sequence 5, Applicatio	1.61e+01
38	89	4.2	1085	1	US-08-485-	Sequence 5, Applicatio	1.61e+01
39	88	4.1	182	1	US-08-127-	Sequence 138, Applicat	1.91e+01
40	88	4.1	182	1	US-08-127-	Sequence 139, Applicat	1.91e+01
41	88	4.1	182	1	US-08-127-	Sequence 171, Applicat	1.91e+01
42	88	4.1	182	1	US-08-127-	Sequence 170, Applicat	1.91e+01
43	88	4.1	535	1	US-08-455-	Sequence 49, Applicati	2.25e+01
44	87	4.1	1088	1	US-08-484-	Sequence 6, Applicatio	2.25e+01
45	87	4.1	1088	1	US-08-484-	Sequence 6, Applicatio	2.25e+01

## ALIGNMENTS

RESULT 1 PCT-US95-13662A-3 STANDARD; PRT; 286 AA.

xxxxxx

Sequence 3, Application PC/TUS9513662A

GENERAL INFORMATION:  
APPLICANT: LENNOX, Tricia L.

APPLICANT: SLATKO, Barton E.

TITLE OF INVENTION: PURIFIED THERMOSTABLE INORGANIC  
OPERATING SYSTEM: PC-DOS/MS-DOS

TITLE OF INVENTION: LITORALIS  
NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESS:  
ADDRESSEE: NEW ENGLAND BIOLABS, INC.

STREET: 32 TOZER ROAD  
CITY: BEVERLY

STATE: MASSACHUSETTS  
COUNTRY: USA

ZIP: 01915

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13662A

FILING DATE:  
CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/329,721

FILING DATE: 25-OCT-1994  
ATTORNEY/AGENT INFORMATION:

NAME: WILLIAMS, GREGORY D.  
REGISTRATION NUMBER: 30901

REFERENCE/DOCKET NUMBER: NEB-105-PCT  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (508) 927-5054  
TELEFAX: (508) 927-1705

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:

LENGTH: 286 amino acids

```

CC      TELEPHONE: (508) 927-5054
CC      TELEFAX: (508) 927-1705
CC      INFORMATION FOR SEQ ID NO: 10:
CC      SEQUENCE CHARACTERISTICS:
CC          LENGTH: 176 amino acids
CC          TYPE: amino acid
CC          TOPOLOGY: linear
CC      MOLECULE TYPE: protein
CC      SEQUENCE 176 AA; 20659 MW; 160121 CN;
SO

Query Match      8.7%; Score 185; DB 2; Length 176;
Best Local Similarity 36.9%; Pred. No. 2,28e-07;
Matches 31; Conservative 20; Mismatches 30; Indels 3; Gaps

Db 66 DDDPDDIVINRREPTVGVLLIARPIGLFKKIDSGDDYKLVANPVDPYFNDKMLISDY 125
OY 116 DNDPDIIVCEISKVCAKEIIGVVLGLANIDGETDMKVIATVDDPDANINDINDV 175
Db 126 -P-KA-FLDEIAHFFORKELOGK 146
OY 176 KRLKPGYLEATVDMFRKRYKVDGR 199

RESULT 3 STANDARD; PRT: 179 AA.
XX ID PC1-US95-13662A-7
XX AC xxxxxx
XX DT
XX DE
XX SEQUENCE 7, Application PC/TUS9513662A
CC GENERAL INFORMATION:
CC APPLICANT: LENNOX, Tricia L.
CC APPLICANT: SIATKO, Barton E.
CC APPLICANT: SEARS, Lauren E.
CC TITLE OF INVENTION: PURIFIED THERMOSTABLE INORGANIC
CC TITLE OF INVENTION: PYROPHOSPHATASE OBTAINABLE FROM THERMOCOCCUS
CC NUMBER OF SEQUENCES: 28
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: NEW ENGLAND BIOLABS, INC.
CC STREET: 32 TOZER ROAD
CC CITY: BEVERLY
CC STATE: MASSACHUSETTS
CC COUNTRY: USA
CC ZIP: 01915
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/13662A
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/329,721
CC FILING DATE: 25-OCT-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: WILKINS, GREGORY D.
CC REGISTRATION NUMBER: 30901
CC REFERENCE/DOCKET NUMBER: NEB-105-PCT
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (508) 927-5054
CC TELEFAX: (508) 927-1705
CC INFORMATION FOR SEQ ID NO: 7:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 179 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: unknown
CC TOPOLOGY: unknown

```



CC MOLECULE TYPE: protein  
SQ SEQUENCE 179 AA; 20465 MW; 177070 CN;  
Query Match 7.7%; Score 165; DB 2; Length 179;  
Best Local Similarity 30.5%; Pred. No. 1.23e-05;  
Matches 32; Conservative 31; Mismatches 38; Indels 4; Gaps 3;  
DB 67 DGDPMADVILISOPTPGCAIKVRPIGMKMNVDGETDNKILAVFDKDPNVS-Y--IKDL 123  
116 DNDPIDVCEISKVCARGEIGVKGILAMIDGETDMKVIAINVDPPDANNDINDV 175  
DB 124 KDVAHLLDELAFNFFSTYKIIIEKKTLYLGWEGKEALKEVSI 168  
OY 176 KRLKPGYLEATVDFWFRYKVPDGKPENEFAFNA-EFKDKFAIDI 219  
RESULT 4  
ID PCT-US95-13662A-6 STANDARD; PRT: 164 AA.  
AC xxxxxx  
DE Sequence 6, Application PC/TUS9513662A  
XX  
CC GENERAL INFORMATION:  
CC APPLICANT: LENNOX, Tricia L.  
CC APPLICANT: SLATKO, Barton E.  
CC APPLICANT: SEARS, Lauren E.  
CC TITLE OF INVENTION: PURIFIED THERMOSTABLE INORGANIC  
CC TITLE OF INVENTION: PYROPHOSPHATASE OBTAINABLE FROM THERMOCOCCUS  
CC NUMBER OF SEQUENCES: 28  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: NEW ENGLAND BIOLABS, INC.  
CC STREET: 32 TOZER ROAD  
CC CITY: BEVERLY  
CC STATE: MASSACHUSETTS  
CC COUNTRY: USA  
CC ZIP: 01915  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent In Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US95/13662A  
CC FILING DATE:  
CC CLASSIFICATION:  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/329,721  
CC FILING DATE: 25-OCT-1994  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: WILLIAMS, GREGORY D.  
CC REGISTRATION NUMBER: 30901  
CC REFERENCE/DOCKET NUMBER: NEB-105-PCT  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (508) 927-5054  
CC TELEFAX: (508) 927-1705  
CC INFORMATION FOR SEQ ID NO: 6:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 164 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: unknown  
CC TOPOLOGY: unknown  
CC MOLECULE TYPE: protein  
SQ SEQUENCE 164 AA; 18792 MW; 129342 CN;  
Query Match 7.2%; Score 154; DB 2; Length 164;  
Best Local Similarity 33.3%; Pred. No. 1.07e-04;  
Matches 28; Conservative 18; Mismatches 35; Indels 3; Gaps 1;

DB 56 DGDPLDLIVITNPPEPCVIDIRVIGYLNWVDSGEEDAKLIGVPEDP---REDEYRSI 112  
116 DNDPIDVCEISKVCARGEIGVKGILAMIDGETDMKVIAINVDPPDANNDINDV 175  
DB 113 EDLPQHLKEIAHFFERYKIDQK 136  
OY 176 KRLKPGYLEATVDFWFRYKVPDGK 199  
RESULT 5  
ID PCT-US95-13662A-4 STANDARD; PRT: 263 AA.  
AC xxxxxx  
DE Sequence 4, Application PC/TUS9513662A  
XX  
CC GENERAL INFORMATION:  
CC APPLICANT: LENNOX, Tricia L.  
CC APPLICANT: SLATKO, Barton E.  
CC APPLICANT: SEARS, Lauren E.  
CC TITLE OF INVENTION: PURIFIED THERMOSTABLE INORGANIC  
CC TITLE OF INVENTION: PYROPHOSPHATASE OBTAINABLE FROM THERMOCOCCUS  
CC NUMBER OF SEQUENCES: 28  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: NEW ENGLAND BIOLABS, INC.  
CC STREET: 32 TOZER ROAD  
CC CITY: BEVERLY  
CC STATE: MASSACHUSETTS  
CC COUNTRY: USA  
CC ZIP: 01915  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent In Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US95/13662A  
CC FILING DATE:  
CC CLASSIFICATION:  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/329,721  
CC FILING DATE: 25-OCT-1994  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: WILLIAMS, GREGORY D.  
CC REGISTRATION NUMBER: 30901  
CC REFERENCE/DOCKET NUMBER: NEB-105-PCT  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (508) 927-5054  
CC TELEFAX: (508) 927-1705  
CC INFORMATION FOR SEQ ID NO: 4:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 263 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: unknown  
CC TOPOLOGY: unknown  
CC MOLECULE TYPE: protein  
SQ SEQUENCE 263 AA; 29712 MW; 339305 CN;  
Query Match 7.1%; Score 151; DB 2; Length 263;  
Best Local Similarity 37.1%; Pred. No. 1.91e-04;  
Matches 23; Conservative 18; Mismatches 21; Indels 0; Gaps 0;  
DB 102 CEDSPMDVILVMEPVLTGSEFLARARATGLPMIDOGKDKITAYACDDPEFRHYRDIK 161  
OY 114 GGDNDPIDVCEISKVCARGEIGVKGILAMIDGETDMKVIAINVDPPDANNDINDV 173  
DB 162 EL 163  
OY 174 DV 175

```

DE XX Sequence 165, Application US/08127954
XX CC Sequence 165, Application US/08127954
CC CC Patent No. 5451512
CC GENERAL INFORMATION:
CC APPLICANT: Apple, Raymond J.
CC APPLICANT: Bugawan, Teodorica L.
CC TITLE OF INVENTION: Methods and Reagents for HLA Class I A
CC TITLE OF INVENTION: Locus DNA Typing
CC NUMBER OF SEQUENCES: 173
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Hoffmann-La Roche Inc.
CC STREET: 340 Kingsland Street
CC CITY: Nutley
CC STATE: New Jersey
CC COUNTRY: U.S.A.
CC ZIP: 07110-1199
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/127,954
CC FILING DATE:
CC CLASSIFICATION: 436
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Petry, Douglas A.
CC REGISTRATION NUMBER: 35,321
CC REFERENCE/DOCKET NUMBER: 8873
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (510) 814-2974
CC TELEFAX: (510) 814-2977
CC INFORMATION FOR SEQ ID NO: 165:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 182 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 162 AA; 21175 MW; 151183 CN;
CC SO
Db 126 LNEDLRSTADMAAQITQR-KWETLAEADQLRA-YIEGTCVEWLRRY-LENGK 176
Oy 147 IDEGETDKVTALINVDPPDANYNIDVVKLKPGYLEAT-VDMFRRTKVPDGK 199
Query Match 4.8%; Score 103; DB 1; Length 182;
Best Local Similarity: 25.9%; Pred. NO.1.45e+00;
Matches 14; Conservative 17; Mismatches 19; Indels 4; Gaps 4
RESULT .8
ID US-08-127-954-146 STANDARD; PRT; 182 AA.
XX AC xxxxxx
XX DT
XX DE
XX Sequence 146, Application US/08127954
CC DE
CC Sequence 146, Application US/08127954
CC Patent No. 5451512
CC GENERAL INFORMATION:
CC APPLICANT: Apple, Raymond J.
CC APPLICANT: Bugawan, Teodorica L.
CC APPLICANT: Erlich, Henry A.
CC TITLE OF INVENTION: Methods and Reagents for HLA Class I A
CC TITLE OF INVENTION: Locus DNA Typing
CC NUMBER OF SEQUENCES: 173
CC CORRESPONDENCE ADDRESS:

```

CC ADDRESSEE: Hoffmann-La Roche Inc.  
CC STREET: 340 Kingsland Street  
CC CITY: Nutley  
CC STATE: New Jersey  
CC COUNTRY: U.S.A.  
CC ZIP: 07110-1199  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patentin Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/127,954  
CC FILING DATE:  
CC CLASSIFICATION: 436  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Petry, Douglas A.  
CC REGISTRATION NUMBER: 35,321  
CC REFERENCE/DOCKET NUMBER: 8873  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (510) 814-2974  
CC TELEFAX: (510) 814-2977  
CC INFORMATION FOR SEQ ID NO: 146:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 182 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: DNA (genomic)  
CC SEQUENCE 182 AA; 21058 MW; 149365 CN;  
  
Query Match 4.7%; Score 101; DB 1; Length 182;  
Best Local Similarity 24.1%; Pred. No. 2,06e+00;  
Matches 13; Conservative 17; Mismatches 20; Indels 4; Gaps 4;  
  
Db 126 LNEIDRSTADMAQITKR-KWEAHEAEQJRA-YLGGTCVEMLRRT-LENGK 176  
Qy 147 IDEGETDKVIAINVDPDPAANYNDINDVYKRLPGYLEAT-VDMPFRRYKVPDGK 199  
  
RESULT 9  
ID US-07-868-353A-14 STANDARD: PRT; 350 AA.  
AC: xxxxxx  
XX  
XX  
DT  
XX  
DE Sequence 14, Application US/07868353A  
XX Sequence 14, Application US/07868353A  
CC Patent No. 5688662  
CC GENERAL INFORMATION:  
CC APPLICANT: Margolskee, Robert F.  
CC TITLE OF INVENTION: Gustducin Materials and Methods  
CC NUMBER OF SEQUENCES: 36  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
CC ADDRESS: Bicknell  
CC STREET: Two First National Plaza, 20 South Clark  
CC CITY: Chicago  
CC STATE: Illinois  
CC COUNTRY: USA  
CC ZIP: 60603  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patentin Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/07/868,353A  
CC FILING DATE: 19920409  
CC CLASSIFICATION: 530

CC ATTORNEY/AGENT INFORMATION:  
CC NAME: No. 5688662and, Greta E.  
CC REGISTRATION NUMBER: P-35,302  
CC REFERENCE/DOCKET NUMBER: 28038/30793  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (312) 346-5750  
CC TELEFAX: (312) 984-9740  
CC INFORMATION FOR SEQ ID NO: 14:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 350 amino acids  
CC TYPE: AMINO ACID  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 350 AA; 39965 MW; 627049 CN;  
  
Query Match 4.5%; Score 95; DB 1; Length 350;  
Best Local Similarity 30.8%; Pred. No. 5.83e+00;  
Matches 20; Conservative 17; Mismatches 23; Indels 5; Gaps 5;  
  
Db 146 DSAGY-LSDERLVTPGYPTEDDVLSRVKT-GILETFSPK-DLNFEMFDVGGRS 202  
Qy 165 DAANTNDINDVYKRL-KPGYLAIVDFR-RYKVPDGKPFNEFAFNAEKDKDFADIIKS 222  
Db 203 ERKKW 207  
Qy 223 THDHW 227  
  
RESULT 10  
ID US-07-868-353A-15 STANDARD: PRT; 354 AA.  
AC: xxxxxx  
XX  
XX  
DT  
XX  
DE Sequence 15, Application US/07868353A  
XX Sequence 15, Application US/07868353A  
CC Patent No. 5688662  
CC GENERAL INFORMATION:  
CC APPLICANT: Margolskee, Robert F.  
CC TITLE OF INVENTION: Gustducin Materials and Methods  
CC NUMBER OF SEQUENCES: 36  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
CC ADDRESS: Bicknell  
CC STREET: Two First National Plaza, 20 South Clark  
CC CITY: Chicago  
CC STATE: Illinois  
CC COUNTRY: USA  
CC ZIP: 60603  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patentin Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/07/868,353A  
CC FILING DATE: 19920409  
CC CLASSIFICATION: 530  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: No. 5688662and, Greta E.  
CC REGISTRATION NUMBER: P-35,302  
CC REFERENCE/DOCKET NUMBER: 28038/30793  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (312) 346-5750  
CC TELEFAX: (312) 984-9740  
CC TELEX: 25-3856  
CC INFORMATION FOR SEQ ID NO: 15:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 354 amino acids

Query Match	4.3%;	Score 92;	DB 1;	Length 182
Best Local Similarity	27.7%;	Pred. No.	9.73e+00;	

XXXXXX  
AC  
XX  
DT  
XX



CC NAME: Petry, Douglas A.  
 CC REGISTRATION NUMBER: 35, 321  
 CC REFERENCE/DOCKET NUMBER: 8873  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (510) 814-2974  
 CC TELEFAX: (510) 814-2977  
 CC INFORMATION FOR SEQ ID NO: 140:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 182 amino acids  
 CC TYPE: amino acid  
 CC STRANDEDNESS: single  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: DNA (genomic)  
 SQ SEQUENCE 182 AA; 21161 MW; 161123 CN;

Query Match 4.3%; Score 92; DB 1; Length 182;  
 Best Local Similarity 27.7%; Pred. NO. 9.73e+00;  
 Matches 13; Conservative 13; Mismatches 17; Indels 4; Gaps 4;

DB 133 WTADMAAQTKH-KWEAAHYAEQLRA-YLEGTCVEMLRV-LENGK 176  
 Oy 154 WKVAINVDDPDPAANIYNDIVKRLKPGYLEAT-VDWFRRYKVPDGR 199

Search completed: Thu Sep 17 16:57:40 1998  
 Job time : 19 secs.

Query Match 48.5%; Score 1035; DB 5; Length 282;

Best Local Similarity 53.0%; Pred. No. 1,56e-166; Matches 151; Conservative 60; Mismatches 64; Indels 10; Gaps 8;

```

Db 2 YTTROIGAKNTLEKYVIEKD-GKPSAFHDIPLYADKNNFNMYVEIPMTNAKLEIT 60
OY 4 FSTERAPAFSLERVFLKNEKQYISPHDIPITADKD--VFHMVVEPRMSNAKMEIA 61
Db 61 KEETLNPIODTKKGLRFVNCFFPHGXIYHNGAFPOTWEDPNVSHETAVGDNNDPID 120
OY 62 TKDPLNPIKQVKKGLRVANLFPYKGYINNGAIPQWEDPNHNRKHTCCGDNNDPID 121
Db 121 VLEIGETIAYGVQKVALGIMALLDEGETDMVYAIADINDPLAPKLNIDIEVEKYFPG 180
OY 122 VCEIGSKVACGEIIGVAVGLIAMDGEEDMWYAIANVDDPRAANYNDI 181
Db 181 LIRATNEFRYKIPDGKPENOFAPSGEAKNKRYALDIKETHDSWKOLJAGKSSDSKGI 240
OY 182 YLEATVDMFRYKVPDGRPENEFAFNAEKDPAIDIKSTHDMKALVT-KKTNGKGI 240
Db 241 DLTNVTLPDP-ty-SKAA-S--DAIPASLAKADAPIDKSIDKWF 280
OY 241 SCANTTISESPFKCDPDAARAIVDALPPEC-ESACTVPTDVDMWF 284

```

## RESULT 4

ENTRY 4 #type complete  
TITLE Inorganic pyrophosphatase (EC 3.6.1.1) product complex, chain A

ALTERNATE\_NAMES  
PDB\_TITLE B - yeast (Saccharomyces cerevisiae)  
ORGANISM prase

STRUCTURE structure of inorganic pyrophosphatase

REFERENCE #formal\_name Saccharomyces cerevisiae  
#note expressed in *Escherichia coli*, strain hb101

REFERENCE #authors Heikinheimo, P.; Lehtonen, J.; Goldman, A.

REFERENCE #submission submitted to the Brookhaven Protein Data Bank, October 1996

REFERENCE #cross-references PDB:1WGB

REFERENCE #authors Heikinheimo, P.; Lehtonen, J.; Baykov, A.; Lahti, R.; Cooperman, B.S.; Goldman, A.

REFERENCE #journal Structure (London) (1996) 4:1491

REFERENCE #title The structural basis for pyrophosphatase catalysis.

REFERENCE #authors Heikinheimo, P.; Salminen, T.; Cooperman, B.S.; Lahti, R.; Goldman, A.

REFERENCE #journal Acta Crystallogr. (1995) D51:399

REFERENCE #title New crystal forms of *e. coli* and *s. cerevisiae* soluble

REFERENCE #comment Resolution: 2.0 angstroms

REFERENCE #comment Determination: X-ray diffraction

REFERENCE #comment R-value: 0.172

REFERENCE #comment hydrolyase; manganese; pyrophosphate phosphohydrolase

REFERENCE #feature 147-149

REFERENCE #feature 165-167

REFERENCE #feature 171-177

REFERENCE #feature 181-191

REFERENCE #feature 194-196

REFERENCE #feature 205-207

REFERENCE #feature 212-231

REFERENCE #feature 235-260

REFERENCE #feature 274-277

REFERENCE #feature 285-267, 2-9, 16-21,

Query Match 48.5%; Score 1035; DB 5; Length 282;

Best Local Similarity 53.0%; Pred. No. 1,56e-166; Matches 151; Conservative 60; Mismatches 64; Indels 10; Gaps 8;

```

Db 2 YTTROIGAKNTLEKYVIEKD-GKPSAFHDIPLYADKNNFNMYVEIPMTNAKLEIT 60
OY 4 FSTERAPAFSLERVFLKNEKQYISPHDIPITADKD--VFHMVVEPRMSNAKMEIA 61
Db 61 KEETLNPIODTKKGLRFVNCFFPHGXIYHNGAFPOTWEDPNVSHETAVGDNNDPID 120
OY 62 TKDPLNPIKQVKKGLRVANLFPYKGYINNGAIPQWEDPNHNRKHTCCGDNNDPID 121
Db 121 VLEIGETIAYGVQKVALGIMALLDEGETDMVYAIADINDPLAPKLNIDIEVEKYFPG 180
OY 122 VCEIGSKVACGEIIGVAVGLIAMDGEEDMWYAIANVDDPRAANYNDI 181
Db 181 LIRATNEFRYKIPDGKPENOFAPSGEAKNKRYALDIKETHDSWKOLJAGKSSDSKGI 240
OY 182 YLEATVDMFRYKVPDGRPENEFAFNAEKDPAIDIKSTHDMKALVT-KKTNGKGI 240
Db 241 DLTNVTLPDP-ty-SKAA-S--DAIPASLAKADAPIDKSIDKWF 280
OY 241 SCANTTISESPFKCDPDAARAIVDALPPEC-ESACTVPTDVDMWF 284

```

## RESULT 5

ENTRY 5 #type complete  
TITLE Inorganic pyrophosphatase (EC 3.6.1.1) metal complex, chain A

ALTERNATE\_NAMES  
PDB\_TITLE B - yeast (Saccharomyces cerevisiae)  
ORGANISM prase

STRUCTURE structure of inorganic pyrophosphatase

REFERENCE #formal\_name Saccharomyces cerevisiae  
#note expressed in *Escherichia coli*, strain hb101

REFERENCE #authors Heikinheimo, P.; Goldman, A.

REFERENCE #submission submitted to the Brookhaven Protein Data Bank, October 1996

REFERENCE #cross-references PDB:1WGB

REFERENCE #authors Heikinheimo, P.; Lehtonen, J.; Baykov, A.; Lahti, R.; Cooperman, B.S.; Goldman, A.

REFERENCE #journal Structure (London) (1996) 4:1491

REFERENCE #title The structural basis for pyrophosphatase catalysis.

REFERENCE #authors Heikinheimo, P.; Salminen, T.; Cooperman, B.S.; Lahti, R.; Goldman, A.

REFERENCE #journal Acta Crystallogr. (1995) D51:399

REFERENCE #title New crystal forms of *e. coli* and *s. cerevisiae* soluble

REFERENCE #comment Resolution: 2.2 angstroms

REFERENCE #comment Determination: X-ray diffraction

REFERENCE #comment R-value: 0.172

REFERENCE #comment hydrolyase; manganese; pyrophosphate phosphohydrolase

REFERENCE #feature 165-167

REFERENCE #feature 171-177

REFERENCE #feature 181-191

REFERENCE #feature 194-196

REFERENCE #feature 204-207

REFERENCE #feature 212-230

REFERENCE #feature 235-260

REFERENCE #feature 274-277

REFERENCE #feature 285-267, 2-9, 16-21,

SUMMARY #length 283 #molecular-weight 31941 #checksum 5403





```

162-165      #region turn (type I)\
178-181      #region turn (type II)\
204-207      #region turn (type II)\
227-240      #region turn (type II)\
247-250      #region turn (type I)\
250-253      #region turn (type I)\
273-276      #region turn (type I)\
273-278      #region turn (type I)\
SUMMARY      #length 282 #molecular-weight 31781 #checksum 830

Query Match      48.4%: Score 1032; DB 5; Length 282;
Best Local Similarity 52.8%: Pred. No. 5.83e-166;
Matches 150; Conservative 58; Mismatches 68; Indels 8; Gaps 7;

Db      2 YTRQIGAKNTLEYKVIEND-GKPSAFHDIPYADKENNIFEMVVEIPMTNAKLEIT 60
OY      4 FSTERRAAPSLEVRFLKNEKGYISPHDIPYADKD--VFHMVVEVRMSAKKEIA 61
Db      61 KEETLNFITDDTAKGKLEFRVNCPEPHGHTHNGAFQPTWEDPNVSHPETRKAGDNDPID 120
OY      62 TKPPLNPKDKVKKGLRYVYANLEFPYKGYIMNYGAIPOQWEDPQHNDRKHGCCDNDPID 121
Db      121 VLEIGETIATGQVKQYKALGIMALLDEGETDMKVIAINDPLAPRLNIDEDYEKTFPG 180
OY      122 VCEIGSVKVCARGELIGVGLIAMDGETDMKVIAINDDDPAANYNDINDVKRLKPG 181
Db      181 LIRATNEMFRIYKIPDGKPENOFAFSGEAKKKKALDIKETHDMSKQLAGKSSDKGI 240
OY      182 YLEATVYMFRRYKYPDKPEREFNFNAEFKDKDAIDIIISTHDHKAALVT-KTNGSKI 240
Db      241 DLATNVLDPDP-TYSKAASDAIPPA-SPKAD-A-PIDKSIDKMF 280
OY      241 SCWMTLSESPFKCDPDAAALVDALPPCESACTVPTDVKMF 284

RESULT 7
ENTRY   #type complete
TITLE   Inorganic pyrophosphatase (EC 3.6.1.1), chain A - yeast
PDB.TITLE
ORGANISM (Saccharomyces cerevisiae)
REFERENCE #formal_name Saccharomyces cerevisiae
AUTHORS A67280
Hartunyanyan, E.H.; Kuranova, I.P.; Lamzin, V.S.; Dauter, Z.;
Wilson, K.S.
#submission submitted to the Brookhaven Protein Data Bank, May 1996
REFERENCE #cross-references PDB:1YPP
AUTHORS TN001408
Hartunyanyan, E.H.; Kuranova, I.P.; Vainshtein, B.K.; Hoehne,
W.E.; Lamzin, V.S.; Dauter, Z.; Teplyakov, A.V.; Wilson,
K.S.
#journal Eur. J. Biochem. (1996) 239:220
#title X-ray structure of yeast inorganic pyrophosphatase complexed
#with manganese and phosphate.
REFERENCE TN001409
Hartunyanyan, E.H.; Oganessyan, V.Yu.; Oganessyan, N.N.;
Terzyan, S.S.; Popov, A.N.; Rubinsky, S.V.; Vainshtein,
B.K.; Nazarova, T.I.; Kurilova, S.A.; Vorobyova, N.N.;
Avaeva, S.M.
#journal Crystallogr. Rep. (Transl. Kristallografiya) (1996) 41:84
#title Structure of inorganic pyrophosphatase from e. coli and its
#complex with a mn2+ ion at 2.2 angstroms resolution.
REFERENCE TN001410
Helkineimo, P.; Salminen, T.; Lahti, R.; Cooperman, B.;
Goldman, A.
#journal Acta Crystallogr. (1995) D51:399
#title New crystal forms of escherichia coli and saccharomyces
#cerevisiae soluble inorganic pyrophosphatases.
REFERENCE TN001411
Teplyakov, A.; Oboolova, G.; Wilson, K.S.; Ishii, K.; Kaji,
H.; Samejima, T.; Kuranova, I.
#journal Protein Sci. (1994) 3:1098
#title Crystal structure of inorganic pyrophosphatase from thermus
#therophilus.

```

---

```

REFERENCE TN001412
AUTHORS Kankare, J.; Neal, G.S.; Salminen, T.; Glunhoff, T.;
#journal Cooperman, B.S.; Lahti, R.; Goldman, A.
#title Protein Eng. (1994) 7:823
#title The structure of e. coli soluble inorganic pyrophosphatase at
#2.7 a resolution.
REFERENCE TN001413
AUTHORS Baykov, A.A.; Shestakov, A.S.
#journal Eur. J. Biochem. (1992) 206:463
#title Two pathways of pyrophosphate hydrolysis and synthesis by
#yeast inorganic pyrophosphatase.
REFERENCE TN001414
Chirgadze, N.Yu.; Kuranova, I.P.; Nevskaya, N.A.; Teplyakov,
A.V.; Wilson, K.S.; Strokopytov, B.N.; Hartunyanyan, E.H.;
Hoehne, W.E.
#journal Kristallografiya (1991) 36:128
#title Crystal structure of mmp1 complex of yeast inorganic
#pyrophosphatase at 2.35 angstroms resolution (Russian).
REFERENCE TN001415
AUTHORS Kuranova, I.P.; Smirnova, E.A.; Chirgadze, N.Yu.
#journal Kristallografiya (1990) 35:1581
#title The growing of crystals of inorganic pyrophosphatase from
#yeast with metal ions and phosphate (Russian).
REFERENCE TN001416
AUTHORS Kuranova, I.P.; Sokolov, V.I.
#journal Bioorg. Khim. (1986) 12:749
#title A conformational hypothesis of the trans-ligation of metals
#which activate pyrophosphatase and related enzymes
#(Russian).
REFERENCE TN001417
Terzyan, S.S.; Voronova, A.A.; Smirnova, E.A.; Kuranova,
I.P.; Nekrasov, Yu.V.; Hartunyanyan, E.G.; Vainshtein, B.K.;
Hoehne, W.; Hansen, G.
#journal Bioorg. Khim. (1984) 10:1469
#title Inorganic pyrophosphatase from yeast at 3 a resolution
#(Russian).
REFERENCE TN001418
Knight, W.B.; Dunaway-mariano, D.; Ransom, S.C.; Villafranca,
J.J.
#journal J. Biol. Chem. (1984) 259:2886
#title Investigations of the metal ion-binding sites of yeast
#inorganic pyrophosphatase.
REFERENCE TN001419
Cooperman, B.S.
#journal Meth. Enzymol. (1982) 87:526
#title The mechanism of action of yeast inorganic pyrophosphatase.
REFERENCE TN001420
Baykov, A.A.; Avaeva, S.M.
#journal Eur. J. Biochem. (1974) 47:57
#title Yeast inorganic pyrophosphatase: studies on metal binding.
COMMENT Resolution: 2.4 angstroms
COMMENT Determination: X-ray diffraction
FEATURES
KEYWORDS hydrolyase
38-40 #region helix (right hand 3-10)\
64-66 #region helix (right hand 3-10)\
165-167 #region helix (right hand 3-10)\
171-177 #region helix (right hand alpha)\
183-191 #region helix (right hand alpha)\
194-196 #region helix (right hand 3-10)\
204-207 #region helix (right hand 3-10)\
212-231 #region helix (right hand alpha)\
150-160,133-140, #region beta barrel\
42-49,92-95, #region beta barrel\
116-123 #region beta barrel\
150-160,133-140, #region beta barrel\
42-49,99-101, #region beta barrel\
116-123 #region beta barrel\
150-160,133-140, #region beta barrel\
42-49,94-96, #region beta barrel\
150-160,141-147, #region beta barrel\
202-204 #region beta barrel\
56-60,67-73,76-76 #region beta barrel\

```

56-60,67-73,16-22, #region beta sheet\  
 1-8,265-267 #region beta sheet\  
 24-26,16-22,1-8, #region beta sheet\  
 82-84,278-280 #region beta sheet\  
 38-41 #region turn (type I)\  
 63-66 #region turn (type II)\  
 159-162 #region turn (type I)\  
 162-165 #region turn (type II)\  
 178-181 #region turn (type II)\  
 204-207 #region turn (type II)\  
 237-240 #region turn (type II)\  
 247-250 #region turn (type I)\  
 250-253 #region turn (type I)\  
 1-276 #region turn (type I)\  
 1-278 #region turn (type I)\  
 #length 282 #molecular-weight 31781 #checksum 830

Query Match 48.4%; Score 1032; DB 5; Length 282;  
 Best Local Similarity 52.8%; Pred. No. 5,836-166; Indels 8; Gaps 7;  
 Matches 150; Conservative 58; Mismatches 68; Indels 8; Gaps 7;

Db 2 YTTTQIGAKNTLEKYKIEKD-GKPSAFHDIPLYADKKNMVEIPMTNAKLEIT 60  
 Oy 4 FSTEEBAAPFSLERYELKNEKQYISPHDIPYADKQ-VFHHVVEYPRMSNAKMEIA 61  
 Db 61 KEETLNPITIDTAKKLRVRCNCFPHGIIHNGAFPOTWEDPNVSHPEKAVGNDPID 120  
 Oy 62 TRDPLNPIDQVAKKLRVYANLEPKYIWNNGALPOTWEDBGHDKGCCGDNDPID 121  
 Db 121 VEIGETIAYTGQVKVAKKLMALDEGETDMKVAIDINDPLAPKLNIDIEDEKYPFG 180  
 Oy 122 VCGISGVKARGELIGVGLIAMIDEGETDMKVAIANVDDPDANYNIDINDVKRLKAG 181  
 Db 181 LRAINEMFRIRYIPDGKPEPNQAFSGEAKNKYALDITKEHDSKOLIAKSSDSKCI 240  
 Oy 182 YLEATVDWFRYKVPDGKPENEFAFNAEKDFALDITKSTHDKALVT-KTINKGI 240  
 Db 241 DLTNVLPDTP-TYSKASDAIPPA-SPKAD-A-PIDKSIDKWF 280  
 Oy 241 SCMTTLESPPKCDPAARAIVDALPPCESACTVPTDVKWF 284

RESULT 8  
 ENTRY 8  
 TITLE PWKL #type complete  
 ORGANISM inorganic pyrophosphatase (EC 3.6.1.1) - yeast (*Kluyveromyces marxianus* var. *lactis*)  
 #formal\_name *Kluyveromyces marxianus* var. *lactis*, *Candida sphaerica*  
 DATE 31-Dec-1991 #sequence-revision 31-Dec-1991 #text-change 05-Sep-1997

ACCESSIONS S07894  
 REFERENCE S07892  
 #authors Stark, M.J.R.; Milner, J.S.  
 #journal Yeast (1989) 5:35-50  
 #title Cloning and analysis of the *Kluyveromyces lactis* TRP1 gene: a chromosomal locus flanked by genes encoding inorganic pyrophosphatase and histone H3.  
 #cross-references W01D:89189093  
 #accession S07894  
 #molecule-type DNA  
 #residues 1-287 #label STA  
 #cross-references EMBL:X14230; NID:62900; FID:62903

GENETICS  
 #gene ITP  
 CLASSIFICATION #superfamily inorganic pyrophosphatase  
 KEYWORDS hydrolyase  
 FEATURE 79,151  
 #active\_site Arg, Glu #status predicted  
 SUMMARY #length 287 #molecular-weight 32165 #checksum 3660

Query Match 47.9%; Score 1022; DB 1; Length 287;  
 Best Local Similarity 51.7%; Pred. No. 4,706-164;

Matches 149; Conservative 65; Mismatches 63; Indels 11; Gaps 10;

Db 1 MS-YTTTQVAKNSLDKYYIEKD-GKPSAFHDIPLYADKKNMVEIPMTNAK 58  
 Oy 1 MSGSTERRAAPFSLERYELKNEKQYISPHDIPYADKQ-VFHHVVEYPRMSNAK 58  
 Db 59 EITKEEPLNPITIDTAKKLRVRCNCFPHGIIHNGAFPOTWEDPNVSHPEKAVGND 118  
 Oy 59 EIAKPDPLNPIDQVAKKLRVYANLEPKYIWNNGALPOTWEDBGHDKGCCGDND 118  
 Db 119 PLDYLEGEVATYGQVKVAKKLMALDEGETDMKVAIDINDPLAPKLNIDIEDEK 178  
 Oy 119 PIDVCEISGVKARGELIGVGLIAMIDEGETDMKVAIANVDDPDANYNIDINDV-KR 177  
 Db 179 L-PELLATNEMFRIRYIPDGKPEPNQAFSGEAKNKYALDITKEHDSKOLIAK 237  
 Oy 178 LKPYLATVDWFRYKVPDGKPENEFAFNAEKDFALDITKSTHDKALVT-KTIN 236  
 Db 238 AKKIDLTNTTSLDTP-TYSKASDAIPPA-SPKAD-A-PIDKSIDKWF 281  
 Oy 237 GKGISCMNTTLESPPKCDPAARAIVDALPPCESACTVPTDVKWF 284

RESULT 9  
 ENTRY 9  
 TITLE ITPP #type complete  
 ORGANISM inorganic pyrophosphatase (EC 3.6.1.1) - yeast (*Saccharomyces cerevisiae*)  
 #formal\_name *Saccharomyces cerevisiae*  
 #authors Harutyunyan, E.H.; Terzyan, S.S.; Vainshtein, B.K.  
 #submission submitted to the Brookhaven Protein Data Bank, February 1983  
 #cross-references PDB:1TPP  
 REFERENCE TNO16962  
 #authors Arutyunyan, E.G.; Terzyan, S.S.; Voronova, A.A.; Kuranova, I.P.; Smirnova, E.A.; Vainshtein, B.K.; Hoehne, W.E.; Hansen, G.  
 #journal Dokl. Akad. Nauk SSSR (1981) 258:1481  
 #title X-ray diffraction study of inorganic pyrophosphatase from baker's yeast at the 3 angstroms resolution (Russian).  
 REFERENCE TNO16963  
 #authors Arutyunyan, E.G.; Terzyan, S.S.; Voronova, A.A.; Kuranova, I.P.; Smirnova, E.A.; Vainshtein, B.K.; Hoehne, W.E.; Hansen, G.  
 #journal Dokl. Biochem. (1981) 258:189  
 #title X-ray structural investigation of inorganic pyrophosphatase from baker's yeast at 3 angstroms resolution.  
 REFERENCE TNO16964  
 #authors Makhdadian, V.V.; Smirnova, E.A.; Voronova, A.A.; Tobis, A.B.; Kuranova, I.P.; Harutyunyan, E.G.; Vainshtein, B.K.; Bienenwald, B.; Hansen, G.; Hoehne, W.E.  
 #journal Kristallografiya (1980) 25:280  
 #title X-ray structural investigation of inorganic pyrophosphatase of baker yeast. II. phase calculation and structure model at 6 angstroms resolution (Russian).  
 REFERENCE TNO16965  
 #authors Makhdadian, V.V.; Smirnova, E.A.; Voronova, A.A.; Tobis, A.B.; Kuranova, I.P.; Arutyunyan, E.G.; Vainshtein, B.K.; Bienenwald, B.; Hansen, G.; Hoehne, W.E.  
 #journal Sov. Phys. Crystallogr. (1980) 25:163  
 #title X-ray structural investigation of inorganic pyrophosphatase of yeasts. II. calculation of phases and structural model at 6 angstroms resolution.  
 REFERENCE TNO16966  
 #authors Smirnova, E.A.; Makhdadian, V.V.; Voronova, A.A.; Kuranova, I.P.; Harutyunyan, E.G.; Vainshtein, B.K.; Heilmann, P.; Hoehne, W.E.  
 #journal Kristallografiya (1980) 25:104  
 #title X-ray structural investigation of inorganic pyrophosphatase of baker yeast. I. Growing of crystals, production of derivatives and determination of the heavy atom positions in them (Russian).  
 REFERENCE TNO16967  
 #authors Smirnova, E.A.; Makhdadian, V.V.; Voronova, A.A.; Kuranova,

```

Yy 4 FSTERBAPPSLEKRVFLKNEKGOYISPFPHDIPYADK-D-VEMHAYEVPRWSNAKKEIA 61
Db 61 KEETLNPDIONTK-GKLRFRVNCPEPHHGYIHTNGAFPOIWEEDNVNHPETRAVGNNDPID 119
Oy 62 TKDELNPDKODVKKGLRLVYANFLPEPYKGYIMNNGALPQIWMEDPQHNDKHKGCCGDNDPID 121
Db 120 VLOIGETIATVGQVKEKVALGINALMDEGETDMKVIAINDNPILAPLNDIEDVEKEFPG 179
Oy 122 VCEIGSRVCAARGELIGKVALIGIAMLIDEGETDMKVIAINVDPPAANYNDINDVYKRLKPG 181
Db 180 LLRLATDEMPRIKYPDCKPENOFSESEAKNKAALDIIEHTNSWQOLAGKSSDSKGI 239
Oy 182 YLEATYVWERIKYKYPDSKPENEFRVNAEFKOKFOAIDIIKSTDHMWALVT-KKTNGKI 240
Db 240 DLIVTLPTDP-TYSKASDAIPPA-SPKAD-A-PIDKSIDKMF 279
Oy 241 SCMTTISESPFKCDPPAARAYDALPPEPSACTVPTDYDKMF 284

RESULT 10
ENTRY #type fragment
TITLE Inorganic pyrophosphatase (EC 3.6.1.1) metal complex, chain
# B. fragment 1 - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES pyrase
PDB_TITLE structure of inorganic pyrophosphatase
ORGANISM #formal_name Saccharomyces cerevisiae
#expressed in Escherichia coli, strain hb101
REFERENCE A68201
#authors Heikinheimo, P.; Goldman, A.
#submission submitted to the Brookhaven Protein Data Bank, October 1996
#cross-references PDB:1WGI
REFERENCE TN001185
#authors Heikinheimo, P.; Lehtonen, J.; Baykov, A.; Lahti, R.;
#Cooperman, B.S.; Goldman, A.
#journal Structure (London) (1996) 4:1491
#title The structural basis for pyrophosphatase catalysis.
REFERENCE TN001186
#authors Heikinheimo, P.; Salminen, T.; Cooperman, B.S.; Lahti, R.;
#Goldman, A.
#journal Acta Crystallogr. (1995) D51:399
#title New crystal forms of e. coli and s. cerevisiae soluble
#inorganic pyrophosphatases.
COMMENT Resolution: 2.2 angstroms
COMMENT Determination: X-ray diffraction
COMMENT R-value: 0.172
KEYWORDS hydrolase; manganese; pyrophosphate phosphohydrolase
FEATURE
171-177 #region helix (right hand alpha)\
181-191 #region helix (right hand alpha)\
194-196 #region helix (right hand 3-10)\
204-207 #region helix (right hand 3-10)\
212-230 #region helix (right hand 3-10)\
2-9,16-21,24-26 #region helix (right hand alpha)\
134-138,42-49, #region beta sheet\
92-96,120-123, #region beta sheet\
150-159,137-147 #region beta sheet\
57-59,68-70 #region beta sheet\
48,56,58,78,89,93, #region beta sheet\
115,117,120,147, #region beta sheet\
152,154,193,193 #region beta sheet\
#site glu, lys, glu, arg, tyr, tyr, asp, asp, asp, asp,
#asp, lys, tyr, lys #label ASB
SUMMARY #length 234 #checksum 1513
Query Match 45.7%; Score 976; DB 5; Length 234;
Best Local Similarity 57.3%; Pred. No. 2,68e-155;
Matches 134; Conservative 48; Mismatches 49; Indels 3; Gaps 2;
Db 2 YTTGQIGAKNTLEKVKYIEKD-GKPVASPHDIPLYAKENKNNFMVYVEIRWTAKLEIT 60
Oy 4 FSTERBAPPSLEKRVFLKNEKGOYISPFPHDIPYADK-D-VEMHAYEVPRWSNAKKEIA 61
Db 61 KEETLNPDIONTKGRKLRFRVNCPEPHHGYIHTNGAFPOIWEEDNVNHPETRAVGNNDPID 120

```

```

OY 62 TKDPLNPKODVKKGLRYVANLEPPYKGYIMNYGALPOTWEDPGHNDKHTGCCGNDNDPID 121
DB 121 VLEIGETIAYTGQVOKVAKALGIMLMDDEGETDMKVIADINDPLAPKLNDEDEVEKYPG 180
OY 122 VCEISKSCARGEIIGVAVLGLAMIDGETDMKVIAINVDDPDPAANNNDINDVKRLKPG 181
DB 181 LIRATNEFRIRKIPDGKPEKNOFAFGSEAKKKKVALDIITKETHDSWKOLINGS 234
OY 182 YLEATVDWFRKRYKVPDGRKPNENFAFNAEFKDKFAIDIIKSTHDMKALVTKKT 235

RESULT 11
ENTRY S11496 #type complete
TITLE Inorganic pyrophosphatase (EC 3.6.1.1) - fission yeast
ALTERNATE_NAMES (Schizosaccharomyces pombe)
ORGANISM pyrophosphate phosphohydrolyase
#formal_name Schizosaccharomyces pombe
DATE 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change
08-Sep-1997

ACCESSIONS S11496
REFERENCE S11496
#authors Kawasaki, I.; Adachi, N.; Ikeda, H.
#journal Nucleic Acids Res. (1990) 18:5888
#title Nucleotide sequence of S. pombe inorganic pyrophosphatase.
#cross-references MUID:91016938
#accession S11496
#molecule_type DNA
#residues 1-289 ##label KAW
#cross-references EMBL:X54301; NID:95013; PID:95014

GENETICS
#gene PPA
#classification #superfamily inorganic pyrophosphatase
KEYWORDS hydrolyase
FEATURE
#length 289 #molecular_weight 32467 #checksum 5076
SUMMARY #active_site Arg #status predicted

Query Match 44.58; Score 950; DB 2; Length 289;
Best Local Similarity 48.68; Pred. No. 2,34e-150;
Matches 140; Conservative 66; Mismatches 72; Indels 10; Gaps 9;

DB 1 MSETYREVALNTLDYOVVEKN-GTPISWHDIPLYANAETILINVAEIPRWTOAKL 59
OY 1 MSGFSTERRAPPSLEIRVFLKNEKGYISPFHDIPYV-DQDVHMYVEVPRMSNAMA 58
DB 60 EITKEATLNDIKODTKKRLFRVNCPPHGGIYIWTGAPQIYEDENVVHPETAKGSD 119
OY 59 EIAIKDPLNDIKODVKKGLRYVANLEPPYKGYIMNYGALPOTWEDPGHNDKHTGCCGND 118
DB 120 PLDVCETGEARGYT-GQVOKVAVLGVALLDEGETDMKVIDVNDPLAPKLNDEDEVER 178
OY 119 PIDVCEIG-SKVCARGEIIGVAVLGLAMIDGETDMKVIAINVDDPDPAANNNDINDVKR 177
DB 179 HMPGLIRATNEFRIRKIPDGKPEKNOFAFGSECKNKRKAEVRECNEMERLITKTDPA 238
OY 178 LKPEYLATVDWFRKRYKVPDGRKPNENFAFNAEFKDKFAIDIIKSTHDMKALVTKKTNG 237
DB 239 KSDSESLVAVSVGSVAN-DPSVSTIPPA-QEL-APA-PVDPSYHKWF 282
OY 238 KG-ISCNMTLTSESFPKCDPAARAIVDALPPCESACTVPTDVDKWF 284

RESULT 12
ENTRY A40867 #type complete
TITLE Inorganic pyrophosphatase (EC 3.6.1.1) precursor,
mitochondrial - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES protein YMR156.09; protein YMR267w
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 27-Mar-1993 #sequence_revision 27-Mar-1992 #text_change
06-Feb-1998
ACCESSIONS A40867
REFERENCE A40867
#authors Lundin, M.; Baltscheffsky, H.; Ronne, H.

```

```

#journal J. Biol. Chem. (1991) 266:12168-12172
#title Yeast PPA2 gene encodes a mitochondrial inorganic
pyrophosphatase that is essential for mitochondrial
function.
#cross-references MUID:91286226
#accession A40867
#molecule_type DNA
#residues 1-310 ##label LUN
#cross-references GB:M81880; NID:q172222; PID:q172223
#accession S54014
REFERENCE S54014
#authors Lye, G.; Churcher, C.M.
#submission submitted to the EMBL Data Library, May 1995
#accession S54479
#molecule_type DNA
#residues 1-310 ##label LYE
#cross-references EMBL:Z49260; NID:q809081; PID:q809090; MIPS:YMR267w
#experimental_source strain AB972

GENETICS
#gene SCD:PPA2; IPP2
#cross-references SGD:S0004880; MIPS:YMR267w
#map_position 13R
#genome nuclear
#classification #superfamily inorganic pyrophosphatase
KEYWORDS hydrolyase; mitochondrion
FEATURE
1-30 #domain transit peptide (mitochondrion) #status
predicted #label TNP\
31-310 #product inorganic pyrophosphatase #status predicted
#label MAY

SUMMARY #length 310 #molecular_weight 35572 #checksum 6304

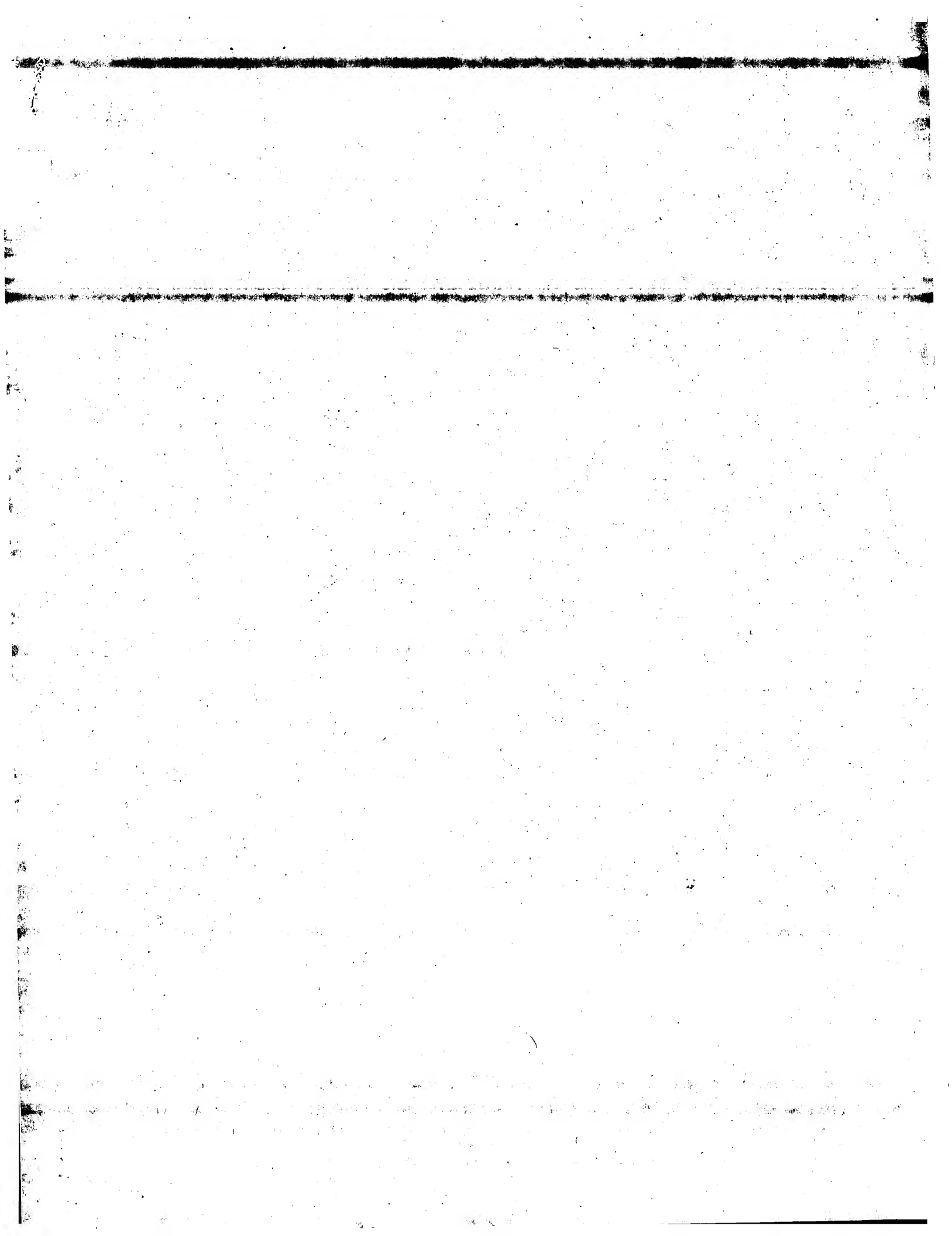
Query Match 38.18; Score 812; DB 2; Length 310;
Best Local Similarity 48.18; Pred. No. 2,99e-124;
Matches 113; Conservative 51; Mismatches 65; Indels 6; Gaps 4;

OY 4 FSTERAAPSLERVFLKNEKGYISPFHDIPYV-YADKD-VFHVAVVEVPRMSNAMAETA 61
DB 94 KELFENPIVDTKNGKRLFRVNNIFPYNGYTHNGALPOTWEDPTIEKLGCVALKGN 153
OY 62 TKDPLNPKODVKKGLRYVANLEPPYKGYIMNYGALPOTWEDPGHNDKHTGC-C--GDN 117
DB 154 DPLDCEISDVLDMGSIKKVYGLSLALIDDELDMKVIADVNDPLSKIDLEKIEE 213
OY 118 DPIDVCEIGSVKARGEIIGVAVLGLAMIDGETDMKVIAINVDDPDPAANNNDINDVKR 177
DB 214 YFPGILDTTRKFRKRYKVPAGKPLNSAFHEQYONSKTITQTKCKNSKNLIS 268
OY 178 LKPYLEATVDWFRKRYKVPDGRKPNENFAFNAEFKDKFAIDIIKSTHDMKALVT 232

RESULT 13
ENTRY S29313 #type complete
TITLE Inorganic pyrophosphatase (EC 3.6.1.1) - Thermoplasma
acidophilum
ORGANISM #formal_name Thermoplasma acidophilum
DATE 22-Jan-1994 #sequence_revision 10-Nov-1995 #text_change
09-Sep-1997
ACCESSIONS S29313
REFERENCE S29313
#authors Richter, O.M.H.; Schaefer, G.
#journal Eur. J. Biochem. (1992) 209:351-355
#title Cloning and sequencing of the gene for the cytoplasmic
inorganic pyrophosphatase from the thermophilic
archaeobacterium Thermoplasma acidophilum.
#accession S29313
#molecule_type DNA
#residues 1-179 ##label RIC
#cross-references EMBL:X64200; NID:q48081; PID:q48082
KEYWORDS hydrolyase
SUMMARY #length 179 #molecular_weight 20465 #checksum 2929

```





\*\*\*\*\*  
 MIPERCH\_PP protein - protein database search, using Smith-Waterman algorithm  
 Run on: Thu Sep 17 16:53:45 1998; Maspar time 10.59 Seconds  
 Tabular output not generated.  
 \*\*\*\*\*

Release 3.1a John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MIPERCH\_PP protein - protein database search, using Smith-Waterman algorithm  
 Run on: Thu Sep 17 16:53:45 1998; Maspar time 10.59 Seconds  
 Tabular output not generated.

Title: >US-08-741-437-1  
 Description: (1-289) from US08741437.pep  
 Perfect Score: 2134  
 Sequence: 1 MSGFSTERAPFSLERVF.....CESACTVPTDVKMFHOKN 289

Scoring table: PAM 150  
 Gap 11

Searched: 69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: swiss-prot35  
 1:swiss1

Statistics: Mean 47.360; Variance 80.422; scale 0.589

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2063	96.7	289	1	IPYR_BOVIN INORGANIC PYROPHOSPHAT	0.00e+00
2	1035	48.5	286	1	IPYR_YEAST INORGANIC PYROPHOSPHAT	1.47e-213
3	1019	47.8	286	1	IPYR_KLUDA INORGANIC PYROPHOSPHAT	1.19e-209
4	939	44.0	288	1	IPYR_SCHPO INORGANIC PYROPHOSPHAT	3.83e-190
5	812	38.1	310	1	IPY2_YEAST INORGANIC PYROPHOSPHAT	2.35e-159
6	658	30.8	114	1	IPYR_HUMAN INORGANIC PYROPHOSPHAT	2.02e-122
7	185	8.7	176	1	IPYR_THEMI THERMOSTABLE INORGANIC	3.54e-16
8	165	7.7	179	1	IPYR_THECI INORGANIC PYROPHOSPHAT	2.21e-12
9	154	7.2	164	1	IPYR_BACP3 INORGANIC PYROPHOSPHAT	2.31e-10
10	151	7.1	263	1	IPYR_ARATH SOLUBLE INORGANIC PYRO	8.03e-09
11	149	7.0	173	1	IPYR_SULAC INORGANIC PYROPHOSPHAT	1.83e-09
12	147	6.9	177	1	IPYR_BABA INORGANIC PYROPHOSPHAT	4.16e-09
13	147	6.9	211	1	IPYR_SOLTU SOLUBLE INORGANIC PYRO	4.16e-09
14	125	5.9	184	1	IPYR_MYCPN INORGANIC PYROPHOSPHAT	2.43e-05
15	122	5.7	176	1	IPYR_HAELN PROBABLE INORGANIC PYR	7.51e-05
16	119	5.6	173	1	IPYR_HELPY INORGANIC PYROPHOSPHAT	4.74e-04
17	117	5.5	169	1	IPYR_SYNZ3 INORGANIC PYROPHOSPHAT	3.29e-04
18	118	5.5	175	1	IPYR_ECOLI INORGANIC PYROPHOSPHAT	4.74e-04
19	117	5.5	184	1	IPYR_MYGE INORGANIC PYROPHOSPHAT	2.86e-03
20	112	5.2	174	1	IPYR_THERP INORGANIC PYROPHOSPHAT	5.79e-03
21	110	5.0	1102	1	YE20_METJA HYPOTHETICAL PROTEIN M	1.64e-02
22	107	4.9	814	1	CADF_HUMAN MUSCLE-CADHERIN PRECUR	4.56e-02
23	104	4.9	365	1	LA03_GORGO CLASS I HISTOCOMPATIBI	

24	103	4.8	362	1	1B56_HUMAN HLA CLASS I HISTOCOMPA	6.38e-02
25	103	4.8	3124	1	CA1C_CHICK COLLAGEN ALPHA 1(XII)	6.38e-02
26	100	4.7	362	1	1B03_GORGO CLASS I HISTOCOMPATIBI	1.72e-01
27	100	4.7	362	1	1B01_GORGO CLASS I HISTOCOMPATIBI	1.72e-01
28	100	4.7	362	1	1B02_GORGO CLASS I HISTOCOMPATIBI	1.72e-01
29	101	4.7	365	1	1A01_SAGOE CLASS I HISTOCOMPATIBI	1.72e-01
30	101	4.7	365	1	1B01_SAGOE CLASS I HISTOCOMPATIBI	1.72e-01
31	101	4.7	370	1	1A03_HUMAN HLA CLASS I HISTOCOMPA	1.24e-01
32	100	4.7	716	1	1P2_FELCA ZONA PELLUCIDA SPERM-B	1.24e-01
33	101	4.7	788	1	CY14_NEUCR SULFATE PERMEASE II	1.24e-01
34	99	4.6	365	1	1A04_PANTR CHIA CLASS I HISTOCOMPA	2.39e-01
35	99	4.6	366	1	1C18_HUMAN HLA CLASS I HISTOCOMPA	2.39e-01
36	99	4.6	369	1	1A1B_MOUSE H-2 CLASS I HISTOCOMPA	2.39e-01
37	98	4.6	730	1	CADF_MOUSE MUSCLE-CADHERIN PRECUR	3.30e-01
38	98	4.6	1750	1	Y832_METJA GENOME POLYPEPTIDE (CO	4.55e-01
39	99	4.6	3430	1	POLG_MNY CLASS I HISTOCOMPATIBI	4.55e-01
40	97	4.5	345	1	HA1F_CHICK GUANINE NUCLEOTIDE-BIN	8.59e-01
41	95	4.5	349	1	GBT1_HUMAN HLA CLASS I HISTOCOMPA	6.26e-01
42	96	4.5	358	1	HLAE_HUMAN HLA CLASS I HISTOCOMPA	6.26e-01
43	96	4.5	362	1	1B19_HUMAN HLA CLASS I HISTOCOMPA	8.59e-01
44	95	4.5	362	1	1B48_HUMAN HLA CLASS I HISTOCOMPA	8.59e-01
45	95	4.5	385	1	YPSC_BACSU HYPOTHETICAL 43.5 KD P	

## ALIGNMENTS

RESULT 1  
 ID IPYR\_BOVIN STANDARD; PRT; 289 AA.

AC P37980;  
 DT 01-OCT-1994 (REL. 30, CREATED)  
 DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE INORGANIC PYROPHOSPHATASE (EC 3.6.1.1) (PYROPHOSPHATE PHOSPHO-  
 DE HYDROLASE) (PAPSE).  
 GN PP.  
 OS BOS TAURUS (BOVINE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 RN EUTHERIA; ARTIODACTYLA.  
 [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE-RETINA;  
 RX MEDLINE; 93077559.  
 RA YANG Z., WENSEL T.G.;  
 RJ J. BIOL. CHEM. 267:24641-24647(1992).  
 CC -1- CATALYTIC ACTIVITY: PYROPHOSPHATE + H(2)O = 2 ORTHOPHOSPHATE.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- TISSUE SPECIFICITY: HIGHEST LEVELS ARE FOUND IN RETINAL ROD OUTER  
 CC SEGMENTS.  
 CC -1- PTM: THE N-TERMINAL IS BLOCKED.  
 CC -1- SIMILARITY: TO OTHER PAPSES.

DR EMBL; M95283; .. NOT\_ANNOTATED\_CDS.  
 DR PIR; A45153; A45153.  
 DR HSSP; P00817; IEPY.  
 DR PROSITE; PS00387; PPAE; 1.  
 KM HYDROLASE; MAGNESIUM.  
 FT ACT\_SITE 57 BY SIMILARITY.  
 SQ SEQUENCE 289 AA; 32844 MW; A1D512C4 CRC32.

Query Match 96.7%; Score 2063; DB 1; Length 289;  
 Best local Similarity 94.8%; Pred. No. 0.00e+00;  
 Matches 214; Conservative 11; Mismatches 4; Indels 0; Gaps 0;

Db 1 MSFSSSERAPFTELEVFLKNEKGYISPFHDIPYADKEVFMVVEPRMSNAKEI 60  
 1 MSGFSTERAPFSLERVFLEKNEKGYISPFHDIPYADKEVFMVVEPRMSNAKEI 60  
 Db 61 ATPKPLPIODVKKGLRVANFPYKGYIWMGALPQWEDSGHDKRTGCCGNDPI 120  
 61 ATPKPLPIODVKKGLRVANFPYKGYIWMGALPQWEDSGHDKRTGCCGNDPI 120  
 Db 121 DVCIGKVCARGEIIVKVLGLAMIDEGETDMKVAIVNVEDPDANVNDINDVRLKP 180  
 121 DVCIGKVCARGEIIVKVLGLAMIDEGETDMKVAIVNVEDPDANVNDINDVRLKP 180

QY 121 DVCISGKVCARGELIGVYGLIAMIDEGEETMVKIAINWDDPDAAINYNDVYKRLKP 180  
 DB 181 GYLEATVDFWERRRYVDPGKPEBNEFAEAEKKNFAIDIESHDYRALVTKTGGKI 240  
 QY 181 GYLEATVDFWERRRYVDPGKPEBNEFAEAEKKNFAIDIESHDYRALVTKTGGKI 240  
 DB 241 SCMMTTSSESPFOCDPDAAKIVDALPPCESACTIPTDVKMFHOKN 289  
 QY 241 SCMMTTSSESPFOCDPDAAKIVDALPPCESACTIPTDVKMFHOKN 289  
 RESULT 2 STANDARD: PRT: 286 AA.  
 ID IPYR YEAST  
 AC P00817;  
 DT 21-JUL-1986 (REL. 01, CREATED)  
 DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE INORGANIC PYROPHOSPHATASE (EC 3.6.1.1) (PYROPHOSPHATE PHOSPHO-  
 HYDROLASE) (PPASE).  
 GN IPPI OR PPAL OR PPA OR YBR011C OR YBR0202.  
 OS SACCAROMYCES CEREVISIAE (BAKER'S YEAST).  
 OC EDUARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMETES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-X2180;  
 RX MEDLINE: 89083474.  
 RA KOLAKOWSKI L.F. JR., SCHLOSSER M., COOPERMAN B.S.;  
 RL NUCLEIC ACIDS RES. 16:10441-10452(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C;  
 RA BAUR A., BOLIS E., MIOSGA T., SCHAFF-GERSTENSCHLAGER I.,  
 RA ZIMMERMANN F.K.;  
 RL SUBMITTED (AUG-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [3]  
 RP SEQUENCE.  
 RX MEDLINE: 78087552.  
 RA COHEN S.A., STERNER R., KEIM P.S., HEINRIKSON R.L.;  
 RL J. BIOL. CHEM. 253:889-897(1978).  
 RN [4]  
 RP SEQUENCE OF 25-35 AND 239-251.  
 RC STRAIN-S288C;  
 RX MEDLINE: 95203288.  
 RA GARELS J.I., FUTCHER B., KOBAYASHI R., LATTER G.I., SCHWENDER B.,  
 RA VOLPE T., WARNER J.R., MCLAUGHLIN C.S.;  
 RL ELECTROPHORESIS 15:1466-1486(1994).  
 RN [5]  
 RP SEQUENCE OF 239-249.  
 RC STRAIN-ATCC 38531 / Y41;  
 RX MEDLINE: 97089742.  
 RA NORECK J., BLUMBERG A.;  
 RL FEMS MICROBIOL. LETT. 137:1-8(1996).  
 RN [6]  
 RP ACTIVE SITE.  
 RA BOND M.W., CHIU N.Y., COOPERMAN B.S.;  
 RL BIOCHEMISTRY 19:94-102(1980).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).  
 RA ARUTUNIAN E.G., TERZIAN S.S., VORONOVA A.A., KURANOVA I.P.,  
 RA SMIRNOVA E.A., VAINSTEIN B.K., HORNE W.E., HANSEN G.;  
 RL DOKL. AKADE. NAUK SSSR 258:1481(1981).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RX MEDLINE: 97148342.  
 RA HEIKINHEIMO P., LEHTONEN J., BAIKOV A., LAHTI R., COOPERMAN B.S.,  
 RL GOLDMAN A.;  
 RL STRUCTURE 4:1491-1508(1996).  
 RN [9]  
 RP SIMILARITY TO E. COLI AND K. LACTIS PPASES.  
 RX MEDLINE: 90254161.  
 RA LAHTI R., KOLAKOWSKI L.F. JR., HEINONEN J., VIHINEN M., POHJANOKSA K.,  
 RA COOPERMAN B.S.;

RL BIOCHIM. BIOPHYS. ACTA 1038:338-345(1990).  
 CC -1- CATALYTIC ACTIVITY: PYROPHOSPHATE + H(2)O = 2 ORTHOPHOSPHATE.  
 CC -1- COFACTOR: THIS ENZYME REQUIRES THE PRESENCE OF DIVALENT METAL  
 CC CATION. MAGNESIUM CONFERS THE HIGHEST ACTIVITY.  
 CC IT BINDS UP TO 4 DIVALENT CATIONS PER SUBUNIT, WITH THREE REQUIRED  
 CC FOR ACTIVITY.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: TO OTHER PPASES.  
 DR EMBL: X13253; G4199;  
 DR EMBL: Z35880; G536206;  
 DR PIR: S45864; PMB;  
 DR PDB: 1PYP; 15-OCT-91.  
 DR PDB: 1YPP; 07-DEC-96.  
 DR PDB: 1MGJ; 19-NOV-97.  
 DR PDB: 1MGJ; 19-NOV-97.  
 DR SWISS-2DPAGE: P00817; YEAST.  
 DR YEPD; 7305;  
 DR SGD: 10000872; IPPI.  
 DR PROSITE: PS00387; PPASE; 1.  
 DR KW HYDROLASE; MAGNESIUM; 3D-STRUCTURE.  
 FT INIT\_MER 0  
 FT ACT\_SITE 56  
 FT BINDING 78  
 FT CONFLICT 40  
 FT CONFLICT 71  
 FT CONFLICT 74  
 FT CONFLICT 123  
 FT CONFLICT 136  
 FT CONFLICT 186  
 FT CONFLICT 224  
 FT CONFLICT 266  
 FT STRAND 4  
 FT STRAND 16  
 FT STRAND 25  
 FT TURN 28  
 FT TURN 38  
 FT STRAND 45  
 FT STRAND 55  
 FT STRAND 79  
 FT TURN 97  
 FT TURN 111  
 FT STRAND 121  
 FT TURN 131  
 FT STRAND 135  
 FT STRAND 138  
 FT STRAND 151  
 FT TURN 160  
 FT TURN 165  
 FT HELIX 172  
 FT TURN 176  
 FT TURN 179  
 FT HELIX 182  
 FT STRAND 203  
 FT HELIX 207  
 FT STRAND 210  
 FT HELIX 212  
 FT TURN 231  
 FT TURN 245  
 FT TURN 251  
 FT TURN 255  
 FT TURN 274  
 SQ SEQUENCE 286 AA; 32184 MW; 0318010F CRC32;  
 Query Match 48.5%; Score 1035; DB 1; Length 286;  
 Best Local Similarity 53.0%; Pred. No. 147e-213;  
 Matches 151; Conservative 60; Mismatches 64; Indels 10; Gaps 8;  
 DB 2 YTTROIGAKNLEKRYVIEKD-GKPSAFHDIPLYADKNNIFNMVVEIPRTNALT 60  
 QY 4 FSTERAPAFLEKRYVFLKKNKGQYISPFHDIPYADKD-VFHMVVEYPRMSNAKKEIA 61



```

Db 61 KEETLPIIDOTFKKGLREFVRCNCFPHGYYIHNGAFQOTWEDPNVSHPETKANGDNDPID 120
OY 62 TTDPLPIIDQVKKGLRYVANLEPYKGYIMNGALPQWEDGHNHDKGCCGDDNDPID 121
Db 121 VLEIGETIAYTGQVKVKAIGMALDEGETDMKVIAIDINDPLAKLNDIEDVEYFPG 180
OY 122 VGEISGVKARGELIIGVYGLIAMLDEGETDMKVIAINVDPAANYNDINDYKRLKPG 181
Db 181 LRLATNEMERIKYIPGKPNENOFAGSGEAKNNKRYALDIKETHDSKOLJAGKSSDSKI 240
OY 182 YLEATVDWFRYKVPDCKPENEFAFNAEFKDFALDIKSTHDKKALVT-KKTNGKI 240
Db 241 DLTNVLPDTP-TY-SKAA-S--DAIPPSLKDADIDKSIDKWF 280
OY 241 SCNMNTLSESFPKCDPDARAIVDALPPEC-ESACTVPTDVDMKF 284

RESULT 3
ID IPYR_KULDA STANDARD; PRT; 286 AA.
AC P1398;
DT 01-JAN-1990 (REL. 13, CREATED)
DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)
DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE INORGANIC PYROPHOSPHATASE (EC 3.6.1.1) (PYROPHOSPHATE PHOSPHO-
HYDROLASE) (PPEASE).
GN IPE1 OR IPE.
OS KLUYVEROMYCES LACTIS (YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 89189093.
RA STARK M.J.R., MILLER J.S.;
RL YEAST 5:35-50(1989).
RN [2]
RP SIMILARITY TO E.COLI AND YEAST PPASES.
RX MEDLINE: 90254161.
RA LAHTI R., KOLAKOWSKI L.F. JR., HEINONEN J., VIHINEN M., POHJANOKSA K.,
COOPERMAN B.S.;
RL BIOCHIM. BIOPHYS. ACTA 1038:338-345(1990).
CC -1- CATALYTIC ACTIVITY: PYROPHOSPHATE + H(2)O = 2 ORTHOPHOSPHATE.
CC -1- COFACTOR: THIS ENZYME REQUIRES THE PRESENCE OF DIVALENT METAL
CATION. MAGNESIUM CONFERS THE HIGHEST ACTIVITY.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: TO OTHER PPASES.
DR EMBL: X14230; G2903;
DR PIR: S07894; PMPYL.
DR HSSP: P00817; IPEP.
DR PROSITE: P500387; PPASE; 1.
KW HYDROLASE; MAGNESIUM.
FT INIT MET 0 BY SIMILARITY.
FT ACT SITE 56 PROBABLE.
FT BINDING 78 INORGANIC PYROPHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 286 AA; 32034 MW; 74833905 CRC32;

Query Match 47.8%; Score 1019; DB 1; Length 286;
Best Local Similarity 51.4%; Pred. No. 1,19e-209;
Matches 147; Conservative 66; Mismatches 63; Indels 10; Gaps 9;

```

```

OY 180 PGYLEATVDWFRYKVPDCKPENEFAFNAEFKDFALDIKSTHDKKALVT-KKTNGK 238
Db 239 KIDLNTLTLDTRA-TYSAEASAVANVLP-DEP-IDKSIDKWF 280
OY 239 GISCNMNTLSESFPKCDPDARAIVDALPPECESACTVPTDVDMKF 284

RESULT 4
ID IPYR_SCHPO STANDARD; PRT; 288 AA.
AC P19117;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE INORGANIC PYROPHOSPHATASE (EC 3.6.1.1) (PYROPHOSPHATE PHOSPHO-
HYDROLASE) (PPEASE).
GN PPA1 OR PPA.
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 91016938.
RA KAWASAKI I., ADACHI N., IKEDA H.;
RL NUCLEIC ACIDS RES. 18:5888-5888(1990).
RN [2]
RP 3D-STRUCTURE MODELLING.
RX MEDLINE: 92337585.
RA VIHINEN M., LUNDIN M., BALTSCHIEFSKY H.;
RL BIOCHIM. BIOPHYS. RES. COMMUN. 186:122-128(1992).
CC -1- CATALYTIC ACTIVITY: PYROPHOSPHATE + H(2)O = 2 ORTHOPHOSPHATE.
CC -1- COFACTOR: THIS ENZYME REQUIRES THE PRESENCE OF DIVALENT METAL
CATION. MAGNESIUM CONFERS THE HIGHEST ACTIVITY.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: TO OTHER PPASES.
DR EMBL: X54301; G5014;
DR PIR: S11496; S11496.
DR HSSP: P00817; IPEP.
DR PROSITE: P500387; PPASE; 1.
KW HYDROLASE; MAGNESIUM.
FT INIT MET 0 PROBABLE.
FT ACT SITE 57 INORGANIC PYROPHOSPHATE (BY SIMILARITY).
FT BINDING 79 79
SQ SEQUENCE 288 AA; 32336 MW; DEED9596 CRC32;

Query Match 44.0%; Score 939; DB 1; Length 288;
Best Local Similarity 48.4%; Pred. No. 3,83e-190;
Matches 139; Conservative 66; Mismatches 72; Indels 10; Gaps 9;

```

[illegible]

ID IPYR\_THEAC STANDARD; PRT; 179 AA.  
AC P37981;  
DT 01-OCT-1994 (REL. 30, CREATED)  
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
DE INORGANIC PYROPHOSPHATASE (EC 3.6.1.1) (PYROPHOSPHATE PHOSPHO-  
HYDROLASE) (PPASE).  
GN PPA.  
OS THERMOPLASMA ACIDOPHILUM.  
OC ARCHAEABACTERIA; EURYARCHAEOTA; THERMOPLASMALES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-DSM 1728;  
RA MEDLINE; 93011147.  
RA RICHTER O.-M.H.; SCHAEFER G.;  
RL EUR. J. BIOCHEM. 209:351-355(1992).  
RN [2]  
RP SEQUENCE OF 1-22. AND CHARACTERIZATION.  
RX MEDLINE; 93011146.  
RA RICHTER O.-M.H.; SCHAEFER G.;  
RL EUR. J. BIOCHEM. 209:343-349(1992).  
CC -1- CATALYTIC ACTIVITY: PYROPHOSPHATE + H(2)O = 2 ORTHOPHOSPHATE.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -1- SIMILARITY: TO OTHER PPASES.  
DR EMBL: X64200; G48082;  
DR PIR: S29282; S29282.  
DR PIR: S29313; S29313.  
DR HSSP: P17288; 1EIP.  
DR PROSITE: PS00387; PPASE; 1.  
KM HYDROLASE; MAGNESIUM.  
FT ACT\_SITE 31  
SQ SEQUENCE 179 AA; 20465 MW; C07397F6 CRC32;  
BY SIMILARITY.  
Query Match 7.7%; Score 165; DB 1; Length 179;  
Best local Similarity 30.5%; Pred. No. 2,21e-12;  
Matches 32; Conservative 31; Mismatches 38; Indels 4; Gaps 3;  
DB 67 DGDPMVDVLTITPPPCVGVDTIRVIGYLNMDSGEDARKLIGVEDP---RFEVRSI 112  
116 DNDPIDVCEIGSKVCARGEIIGVKVGLIAMIDGSETDMKVIAINVDPPAANYNDINDV 175  
DB 124 KDVAHHLDELAFNEFTYKILEKETKYLGMEGKALKEIEVSI 168  
176 KRLKPGYLEATVDMFRRYKVPDGRKPEMEFARNA-EFKDKDEAIDI 219  
QY  
RESULT 9  
ID IPYR\_BACPA3 STANDARD; PRT; 164 AA.  
AC P19514;  
DT 01-FEB-1991 (REL. 17, CREATED)  
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE INORGANIC PYROPHOSPHATASE (EC 3.6.1.1) (PYROPHOSPHATE PHOSPHO-  
HYDROLASE) (PPASE).  
GN PPA.  
OS BACILLUS PS3 (THERMOPHILIC BACTERIUM PS-3).  
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
RN [1]  
RP SEQUENCE.  
RX MEDLINE: 91154162.  
RA ICHIBA T.; TAKENAKA O.; SAMEJIMA T.; HACHIMORI A.;  
RL J. BIOCHEM. 108:572-578(1990).  
CC -1- CATALYTIC ACTIVITY: PYROPHOSPHATE + H(2)O = 2 ORTHOPHOSPHATE.  
CC -1- COFACTOR: THIS ENZYME REQUIRES THE PRESENCE OF DIVALENT METAL  
CATION. MAGNESIUM CONFERS THE HIGHEST ACTIVITY.  
CC -1- SUBUNIT: HOMOTRIMER. IN PRESENCE OF DIVALENT CATIONS THE TRIMERS  
AGGREGATE TO FORM A HEXAMER.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -1- SIMILARITY: TO OTHER PPASES.  
DR PIR: JX0135; JX0135.  
DR HSSP: P17288; 1EIP.  
DR PROSITE: PS00387; PPASE; 1.  
KM HYDROLASE; MAGNESIUM.

FT ACT\_SITE 20 20 BY SIMILARITY.  
SQ SEQUENCE 164 AA; 18792 MW; FF88D7C4 CRC32;  
Query Match 7.2%; Score 154; DB 1; Length 164;  
Best local Similarity 33.3%; Pred. No. 2,31e-10;  
Matches 28; Conservative 18; Mismatches 35; Indels 3; Gaps 1;  
DB 56 DGDPLDILVITTPPPCCVGVDTIRVIGYLNMDSGEDARKLIGVEDP---RFEVRSI 112  
116 DNDPIDVCEIGSKVCARGEIIGVKVGLIAMIDGSETDMKVIAINVDPPAANYNDINDV 175  
DB 113 EDLPQHKLEIAHFFERYKDLQK 136  
176 KRLKPGYLEATVDMFRRYKVPDGR 199  
QY  
RESULT 10  
ID IPYR\_ARATH STANDARD; PRT; 263 AA.  
AC P21216;  
DT 01-MAR-1992 (REL. 21, CREATED)  
DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE SOLUBLE INORGANIC PYROPHOSPHATASE (EC 3.6.1.1) (PYROPHOSPHATE PHOSPHO-  
HYDROLASE) (PPASE).  
GN PPA.  
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).  
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;  
OC CAPRALES; CRUCIFERAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA;  
RX MEDLINE; 91370878.  
RA KIEBER J.J.; SIGNER E.R.;  
RL PLANT MOL. BIOL. 16:345-348(1991).  
CC -1- CATALYTIC ACTIVITY: PYROPHOSPHATE + H(2)O = 2 ORTHOPHOSPHATE.  
CC -1- COFACTOR: ACTIVITY DEPENDENT ON MG2+ (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).  
CC -1- SIMILARITY: TO OTHER PPASES.  
DR EMBL: X57545; G16348;  
DR PIR: S13379; S13379.  
DR HSSP: P17288; 1EIP.  
DR PROSITE: PS00387; PPASE; 1.  
KM HYDROLASE; MAGNESIUM.  
FT ACT\_SITE 68  
SQ SEQUENCE 263 AA; 29712 MW; F6B4A9E4 CRC32;  
BY SIMILARITY.  
Query Match 7.1%; Score 151; DB 1; Length 263;  
Best local Similarity 37.1%; Pred. No. 8,03e-10;  
Matches 23; Conservative 18; Mismatches 21; Indels 0; Gaps 0;  
DB 102 CGSDPMVDVLTITPPPCVGVDTIRVIGYLNMDSGEDARKLIGVEDP---RFEVRSI 112  
114 GNDPIDVCEIGSKVCARGEIIGVKVGLIAMIDGSETDMKVIAINVDPPAANYNDINDV 173  
DB 162 EL 163  
174 DV 175  
QY  
RESULT 11  
ID IPYR\_SULAC STANDARD; PRT; 173 AA.  
AC P50308;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE INORGANIC PYROPHOSPHATASE (EC 3.6.1.1) (PYROPHOSPHATE PHOSPHO-  
HYDROLASE) (PPASE).  
GN PPA.  
OS SULFOLOBUS ACIDOCALDARIUS.  
OC ARCHAEABACTERIA; CRENARCHAEOTA; SULFOBALES.  
RN [1]  
RP SEQUENCE FROM N.A.; PARTIAL SEQUENCE, AND CHARACTERIZATION.  
RC STRAIN-DSM 639;

RX MEDLINE: 95289717.  
 RA MEYER W., MOLL R., KATH T., SCHAEFER G.;  
 RL ARCH. BIOCHEM. BIOPHYS. 319:149-156(1995).  
 CC -1- CATALYTIC ACTIVITY: PYROPHOSPHATE + H(2)O = 2 ORTHOPHOSPHATE.  
 CC -1- COFACTOR: ABSOLUTE REQUIREMENT FOR MAGNESIUM.  
 CC -1- SUBUNIT: HOMOTETRAMER. CYTOPLASMIC.  
 CC -1- SUBCELLULAR LOCATION: TO OTHER PPASES.  
 DR EMBL: X81842; G886704; -  
 DR PROSITE: PS00387; PPASE; 1.  
 KW HYDROLASE; MAGNESIUM.  
 FT ACT\_SITE 26  
 SQ SEQUENCE 173 AA; 19381 MW; 4FAE97B3 CRC32;  
 BY SIMILARITY.  
 Query Match 7.0%; Score 149; DB 1; Length 173;  
 Best Local Similarity 33.7%; Pred. No. 1,83e-09;  
 Matches 29; Conservative 22; Mismatches 31; Indels 4; Gaps 4;  
 Db 62 DDDPLVLTNTQVLP-GSVIEVRPIGILYMKDEGDAKIVAVPKDTPS-FSNIKD 119  
 QY 116 DNDPIDVCEIGS-KVCAREIIGKVLGILAMIDEGEDMKVIALINVDPPANINDIND 174  
 Db 120 INDLPQATKKNKIVFEEHYKELEPGK 145  
 QY 175 VKRLKPGYLEATVDFRRRYKVPD-GK 199  
 RESULT 12 STANDARD; PRT; 177 AA.  
 ID IPYR BARBA  
 AC P51064;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE INORGANIC PYROPHOSPHATASE (EC 3.6.1.1) (PYROPHOSPHATE PHOSPHO-  
 DE HYDROLASE) (PPASE).  
 GN PPA.  
 OS BARTONELLA BACILLIFORMIS.  
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; RICKETTSIAS; RICKETTSIALES;  
 OC BARTONELLACEAE.  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-KC583;  
 RA MITCHELL S.J., MINNICK M.F.;  
 RL SUBMITTED (AUG-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -1- CATALYTIC ACTIVITY: PYROPHOSPHATE + H(2)O = 2 ORTHOPHOSPHATE.  
 CC -1- COFACTOR: THIS ENZYME REQUIRES THE PRESENCE OF DIVALENT METAL  
 CC -1- CATION. MAGNESIUM CONFERS THE HIGHEST ACTIVITY. BINDS 4 DIVALENT  
 CC CATIONS PER SUBUNIT (BY SIMILARITY).  
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: TO OTHER PPASES.  
 DR EMBL: L46591; G940255; -  
 DR PROSITE: PS00387; PPASE; 1.  
 KW HYDROLASE; MAGNESIUM.  
 FT ACT\_SITE 30  
 SQ SEQUENCE 177 AA; 20064 MW; 31852443 CRC32;  
 BY SIMILARITY.  
 Query Match 6.9%; Score 147; DB 1; Length 177;  
 Best Local Similarity 30.4%; Pred. No. 4.16e-09;  
 Matches 35; Conservative 23; Mismatches 53; Indels 4; Gaps 3;  
 Db 66 DSDPIDVLCNTRPLIPGCVINVRPIGALIMDDGSGKEKIIAVPTPL-NGOYIGIDY 124  
 QY 116 DNDPIDVCEIGSKVCARGEIIIGKVLGILAMIDEGEDMKVIALINVDPPANINDINDV 175  
 Db 125 TDLTENILKKIEHFEKHYK--DLEAGK-WAKIEGWKDKNFARLLIQAIERAKAI 176  
 QY 176 KRLKPGYLEATVDFRRRYKVPDCKPNERFAFAFEDKDFALDIKSTHDKWAL 230  
 RESULT 13 STANDARD; PRT; 211 AA.  
 ID IPYR SOLTU  
 AC Q43187;

DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE SOLUBLE INORGANIC PYROPHOSPHATASE (EC 3.6.1.1) (PYROPHOSPHATE PHOSPHO-  
 DE HYDROLASE) (PPASE).  
 GN PPA.  
 OS SOLANUM TUBEROSUM (POTATO).  
 OC EUKARYOTA; PLANTAE; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;  
 OC SOLANALES; SOLANACEAE.  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 96163189.  
 RA DU JARDIN P., ROJAS-BELTRAN J., GEBHARDT C., BRASSEUR R.;  
 RL PLANT PHYSIOL. 109:853-860(1995).  
 CC -1- CATALYTIC ACTIVITY: PYROPHOSPHATE + H(2)O = 2 ORTHOPHOSPHATE.  
 CC -1- COFACTOR: ACTIVITY DEPENDENT ON MG2+.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: TO OTHER PPASES.  
 DR EMBL: Z36894; G534916; -  
 DR PROSITE: PS00387; PPASE; 1.  
 KW HYDROLASE; MAGNESIUM.  
 FT ACT\_SITE 61  
 SQ SEQUENCE 211 AA; 24261 MW; 8E8FC92C CRC32;  
 BY SIMILARITY.  
 Query Match 6.9%; Score 147; DB 1; Length 211;  
 Best Local Similarity 37.2%; Pred. No. 4.16e-09;  
 Matches 32; Conservative 16; Mismatches 35; Indels 3; Gaps 1;  
 Db 95 CEDNDPMDVILMOEVLDFGCLFARARIGLMPIDGEGDKIIVACADDPYRHYTDI- 153  
 QY 114 GGDNDPIDVCEIGSKVCAREIIGKVLGILAMIDEGEDMKVIALINVDPPANINDIND 173  
 Db 154 --KOLPFRALRIREFEDKKNENK 177  
 QY 174 DVKRLKPGYLEATVDFRRRYKVPDCK 199  
 RESULT 14 STANDARD; PRT; 184 AA.  
 ID IPYR MYCEN  
 AC P75250;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE INORGANIC PYROPHOSPHATASE (EC 3.6.1.1) (PYROPHOSPHATE PHOSPHO-  
 DE HYDROLASE) (PPASE).  
 GN PPA.  
 OS MYCOPLASMA PNEUMONIAE.  
 OC PROKARYOTA; TENERICUTES; MOLICUTES; MYCOPLASMA; MYCOPLASMATALES;  
 OC MYCOPLASMATACEAE.  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 29342 / M129;  
 RA MEDLINE: 97105885.  
 RA HIMMELREICH R., HILBERT H., FLAGENS H., PIRKL E., LI B.-C.;  
 RL NUCLEIC ACIDS RES. 24:4420-4449(1996).  
 CC -1- CATALYTIC ACTIVITY: PYROPHOSPHATE + H(2)O = 2 ORTHOPHOSPHATE.  
 CC -1- COFACTOR: THIS ENZYME REQUIRES THE PRESENCE OF DIVALENT METAL  
 CC -1- CATION. MAGNESIUM CONFERS THE HIGHEST ACTIVITY. BINDS 4 DIVALENT  
 CC CATIONS PER SUBUNIT (BY SIMILARITY).  
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: TO OTHER PPASES.  
 DR EMBL: AE000029; G1673987; -  
 DR PROSITE: PS00387; PPASE; 1.  
 KW HYDROLASE; MAGNESIUM.  
 FT ACT\_SITE 19  
 SQ SEQUENCE 184 AA; 21369 MW; D5F880FD CRC32;  
 BY SIMILARITY.  
 Query Match 5.9%; Score 125; DB 1; Length 184;  
 Best Local Similarity 28.6%; Pred. No. 2.43e-05;  
 Matches 18; Conservative 22; Mismatches 20; Indels 3; Gaps 3;

DB 72 GVVVPTRIYVGALEWVDGELDTKLGV-ID-CDP-RYKEINSVNDLPKRVDEIGFIKT 128  
 OY 133 GEIIVGVKIGILAMIDEGSETDMKVIAINVDDPDANYNINDVKRLKPGYLEATVDFERR 192  
 DB 129 YKL 131  
 OY 193 YKV 195

## RESULT 15

ID IPYR.HAEIN STANDARD: PRT: 176 AA.

AC P44529; 01-NOV-1995 (REL. 32, CREATED)

DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)

DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)

DE PROBABLE INORGANIC PYROPHOSPHATASE (EC 3.6.1.1) (PYROPHOSPHATE

DE PHOSPHO-HYDROLASE) (PPASE).

CN PPA OR HI0124.

OS HAEMOPHILUS INFLUENZAE.

OC PROKARYOTA: GRACILICUTES; SCOTOBACTERIA: FACULTATIVELY ANAEROBIC RODS;

OC PASTEURRELLACEAE.

RN 11

RP SEQUENCE FROM N.A.

RC STRAIN-RD / KW20.

RX MEDLINE: 95350630.

RA FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,

RA KERLAVAGE A.R., BULT C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,

RA MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,

RA SCOTT J.D., SHIREY R., LIU L.-I., GLODER A., KELLEY J.M.,

RA WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLOM E., COTTON M.D.,

RA UTTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDER D.M., BRANDON R.C.,

RA FINE L.D., FITCHMAN J.L., FUHRMANN J.L., GEOGHAGEN N.S.M.,

RA GNEHM C.L., MCDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,

RA VENTER J.C.;

RL SCIENCE 269:496-512(1995).

CC -1- CATALYTIC ACTIVITY: PYROPHOSPHATE + H(2)O = 2 ORTHOPHOSPHATE.

CC -1- COFACTOR: THIS ENZYME REQUIRES THE PRESENCE OF DIVALENT METAL

CC CATION. MAGNESIUM CONFERS THE HIGHEST ACTIVITY. BINDS 4 DIVALENT

CC CATIONS PER SUBUNIT (BY SIMILARITY).

CC -1- SUBUNIT: HOMOHEXAMER (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.

CC -1- SIMILARITY: TO OTHER PPASES.

DR EMBL: U32698; G1573079; -

DR PROSITE: PS00387; PPASE; 1.

DR TIGR: H10124; -

KW HYDROLASE; MAGNESIUM.

FT ACT\_SITE 31

SQ SEQUENCE 176 AA; 19725 MW; CEBBD6F CRC32;

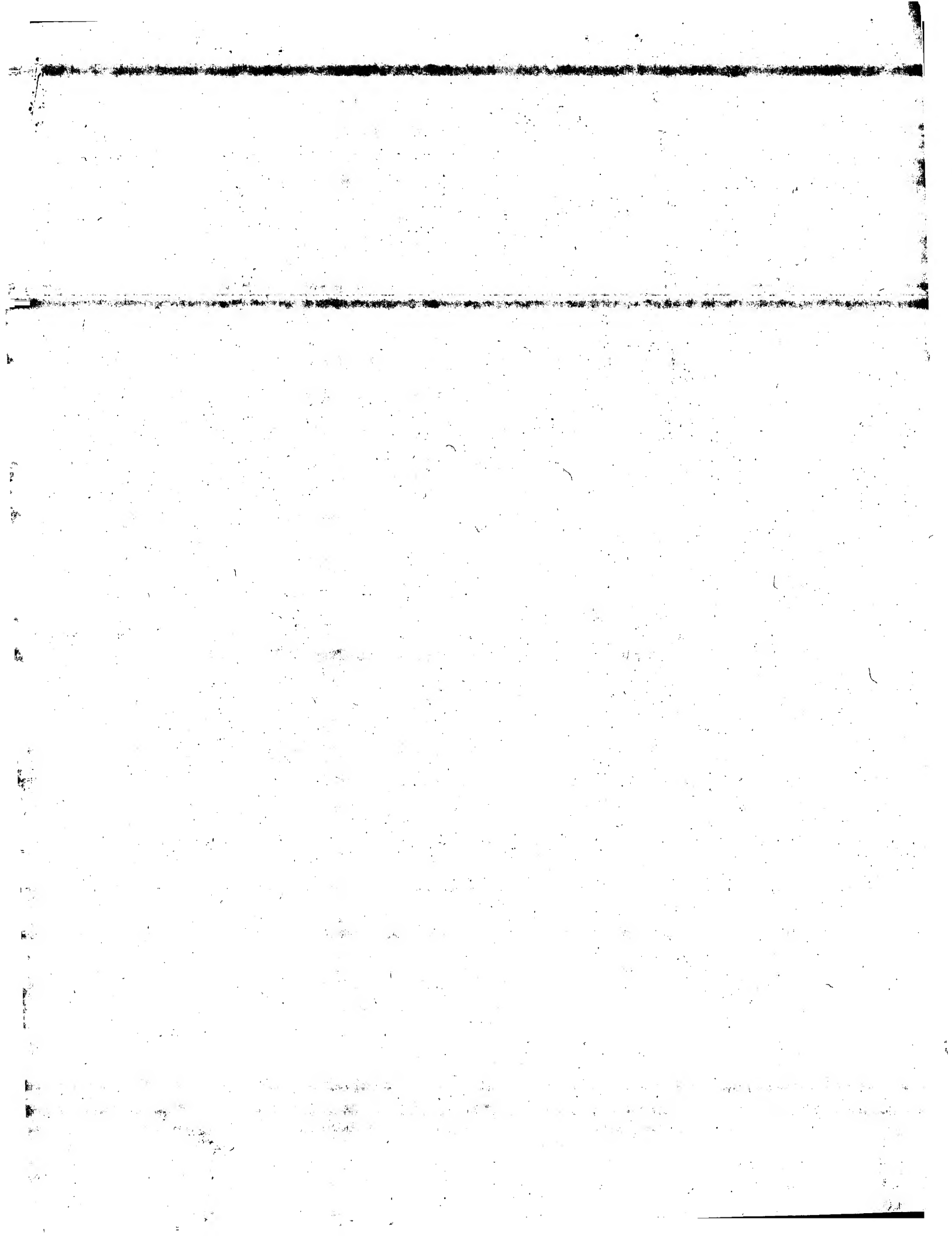
Query Match 5.7%; Score 122; DB 1; Length 176;

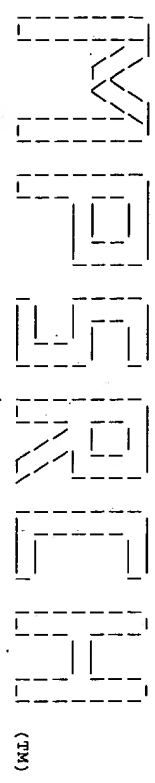
Best Local Similarity 29.6%; Pred. No. 7.51e-05;

Matches 34; Conservative 26; Mismatches 46; Indels 9; Gaps 7;

DB 67 DGEDELVLITRIOPATGVLEAKVIGVKKFEVDGEVDKIVCPADPDRTGNAYNSIAD 126  
 OY 116 DNDPIDVCEIGSKVARGEIIIGVYGLIAMIDEGSETDMKVIAINVDDPDAN-YNDIND 174  
 DB 127 ---L-PAULIKIEFHNNYKALK-KPGST-KV-THMGVDEAKVIRESIKRW 174  
 OY 175 VKRLKPGYLEATVDM-FRRYKVPDGKPNENFAFNAEFKDKDFALIDIKSTHDMK 228

Search completed: Thu Sep 17 16:54:43 1998  
 Job time : 58 secs.





Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Msrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Sep 17 16:55:00 1998; Maspar time 18.79 Seconds  
647.612 Million cell updates/sec

Tabular output not generated.

Title: >US-08-741-437-1  
Description: (1-289) from US08741437.pep  
Perfect Score: 2134  
Sequence: 1 MSGFSTERAPFSLERYRVF.....CESACTVPTVDKMFHHQKN 289

Scoring table: PAM 150  
Gap 11

Searched: 140555 seqs, 42109429 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: sprenb16  
1:sp\_fungi 2:sp\_human 3:sp\_invertebrate 4:sp\_mammal  
5:sp\_mhc 6:sp\_organelle 7:sp\_phage 8:sp\_plant  
9:sp\_bacteria 10:sp\_rodent 11:sp\_virus 12:sp Vertebrate  
13:sp\_unclassified

Statistics: Mean 46.257; Variance 82.300; scale 0.562

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	1119	52.4	292	3	Q18680	HYPOTHETICAL PROTEIN C	6.23e-224
2	1005	47.1	285	1	Q13505	INORGANIC PYROPHOSPHAT	2.86e-197
3	864	40.5	286	1	P87118	HYPOTHETICAL 32.9 KD P	1.72e-164
4	154	7.2	164	9	Q05724	PROPHOSPHATASE (FRAGM	8.78e-10
5	143	6.7	164	9	Q23979	SOLUBLE INORGANIC PYRO	6.87e-08
6	135	6.3	176	9	Q26363	INORGANIC PYROPHOSPHAT	1.50e-06
7	126	5.9	136	9	Q49071	INORGANIC PYROPHOSPHAT	4.35e-05
8	123	5.8	214	8	Q22537	INORGANIC PYROPHOSPHAT	1.30e-04
9	117	5.5	178	9	Q34955	INORGANIC PYROPHOSPHAT	1.12e-03
10	108	5.1	363	5	Q30223	MHC CLASS I ATBE-G*03	2.50e-02
11	109	5.1	482	8	Q41676	LEGUMIN A PRECURSOR.	1.78e-02
12	106	5.0	356	5	Q30444	MHC CLASS I CAVA-G*04	4.87e-02
13	107	5.0	357	5	Q30914	MHC CLASS I A (FRAGMEN	3.49e-02
14	107	5.0	367	5	Q08680	RT1 CLASS I HISTOCOMPA	3.49e-02
15	107	5.0	377	5	Q19445	MHC CLASS IB ANTIGEN.	3.49e-02
16	104	4.9	181	5	Q05409	MHC CLASS I A ANTIGEN	9.45e-02
17	104	4.9	358	5	Q30593	MHC CLASS I ANTIGEN MA	9.45e-02
18	102	4.8	91	5	Q05555	(STRAIN BM6) ALPHA-2 D	1.81e-01
19	103	4.8	181	5	Q29766	MHC TYPE I ANTIGEN (FR	1.31e-01
20	102	4.8	356	5	Q30445	MHC CLASS I CAVA-G*05	1.81e-01

21	103	4.8	362	5	P79523	MHC CLASS I HISTOCOMPA	1.31e-01
22	103	4.8	1462	9	Q34313	YEKN PROTEIN.	1.31e-01
23	101	4.7	166	5	Q95518	HISTOCOMPATIBILITY ANT	2.50e-01
24	101	4.7	166	5	Q95516	HISTOCOMPATIBILITY ANT	2.50e-01
25	100	4.7	181	5	Q19354	MHC CLASS I A ANTIGEN	3.45e-01
26	101	4.7	294	5	Q30714	MHC CLASS I ANTIGEN MA	2.50e-01
27	100	4.7	344	5	P79602	MATURE ALPHA CHAIN OF	3.45e-01
28	100	4.7	354	5	Q95514	HISTOCOMPATIBILITY ANT	3.45e-01
29	100	4.7	354	5	Q95513	HISTOCOMPATIBILITY ANT	3.45e-01
30	101	4.7	363	5	Q19642	MHC CLASS I ANTIGEN HL	2.50e-01
31	101	4.7	363	5	Q29840	MHC HLA-A*0301 BLANK G	2.50e-01
32	101	4.7	365	5	Q19756	MHC CLASS I HLA-A.	2.50e-01
33	101	4.7	3063	2	Q99715	COLLAGEN TYPE XII ALPH	2.50e-01
34	99	4.6	91	5	Q31189	MHC CLASS I H2-K-ALPHA	4.74e-01
35	99	4.6	91	5	P79617	MHC CLASS I HLA-C (FRA	4.74e-01
36	99	4.6	91	5	Q19635	MHC CLASS I ANTIGEN HL	4.74e-01
37	99	4.6	181	5	Q19505	HLA-CW6 (FRAGMENT).	4.74e-01
38	98	4.6	246	5	Q29945	MHC CLASS I HLA-A CELL	6.49e-01
39	99	4.6	347	4	Q29505	GALACTOSIDE 2-L-FUCOSY	4.74e-01
40	98	4.6	356	5	Q30171	LYMPHOCYTE ANTIGEN.	6.49e-01
41	99	4.6	357	5	Q30917	MHC CLASS I A (FRAGMEN	4.74e-01
42	98	4.6	358	5	Q30291	MHC CLASS I PRECURSOR.	6.49e-01
43	99	4.6	366	5	Q29991	HLA CLASS I HEAVY CHAI	4.74e-01
44	99	4.6	366	5	Q31605	ALPHA CHAIN OF MHC CLA	4.74e-01
45	99	4.6	366	5	Q29864	HLA-C ALPHA CHAIN PREC	4.74e-01

ALIGNMENTS

RESULT 1  
ID Q18680; PRELIMINARY; PRT; 292 AA.  
AC Q18680;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL PROTEIN C47E12.4.  
GN C47E12.4.  
OS CAENORHABDITIS ELEGANS.  
OC EUKARYOTA; METAZOA; ACCELLOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA COLES L.  
RL SUBMITTED (JAN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA MEDLINE: 94150718.  
RX WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J.,  
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULLON L.,  
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIJFEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,  
RL NATURE 368:32-38(1994).  
DR EMBL: Z68882; E348322;  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 292 AA; 32363 MW; A31DC923 CRC32;

Query Match 52.4%; Score 1119; DB 3; Length 292;  
Best Local Similarity 53.2%; Pred. No. 6.23e-224;  
Matches 150; Conservatly 65; Mismatches 60; Indels 7; Gaps 6;

DB	11	YEAVERGSLSYSDRYVYIGPOG-IYSPHDIPLFANKDKRYNMIVETIPRTNKMENA	69
QY	4	FSTERRAAPFSLERYFLNKGQVISPFDIPIYADKO--VFHWVVEPRKSNMAMEIA	61
DB	70	TREPSPIKODEKGVAVVNIIFPKKGIINYGALPQTMEDPNHVPDTGAKGNDPID	129
QY	62	TDDPMPIDQVKKGLRYVAMLPFKGIINYGALPQTMEDPNHVPDTGAKGNDPID	121

Db 130 VIEWSGVARGAVLYGTALIDEGEDTMDKLVADVDNDENADKLNDIDVEKYBPG 189  
 122 VCEIGSKVARGELIGVGLIAMDDEGETMDKLVAINVDDPDANNDINDVYKRLPG 181  
 Db 190 LLAASWEFNNYIPAGKPAEFAENGKEKREYAEKVDELNEVWKLII-KEANP-SLN 247  
 182 YLEATYDMEFRYKVPDCKPENEFANAEFKDFAIDIIKSTHDMKALVYTKTKTGKIS 241  
 Db 248 TV-SRVEAVHOGTDEAATAIGATPEHGANA-PLPGVDYDKW 287  
 242 CMNTLSESPKCDPDARAIVALLPPCHSACTVPTVDYDKW 283

RESULT 2 PRELIMINARY: PRT: 285 AA.  
 ID 013505;  
 AC 013505;  
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)  
 DE INORGANIC PYROPHOSPHATASE (PYROPHOSPHATE PHOSPHO-HYDROLASE) (EC 3.6.1.1).  
 GN IPI.  
 OS PICHIA PASTORIS (YEAST).  
 OC PLASMID PEST316-GAL1.  
 CC EURARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCYCETES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-NREL.Y-11430.  
 RA COSANO I., ALVAREZ P., MOLINA M., NOMBELA C.;  
 RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL; AJ001000; E1180018;  
 KW HYDROLASE; PLASMID.  
 SQ SEQUENCE 285 AA; 32068 MW; 60196407 CRC32;

Query Match 47.1%; Score 1005; DB 1; Length 285;  
 Best Local Similarity 52.4%; Pred. No. 2,86e-197;  
 Matches 150; Conservative 56; Mismatches 71; Indels 9; Gaps 8;

Db 1 MS-YSTROIGANTLENVFIKED-GQVVSPPHDIPLADSEKYLNVVVPVPTNAKL 58  
 1 MSGFSTEEBAPPSLEYRFLKNEKQVISPFDIPYAD-KDVEHNVVEPRSNKAK 58  
 Db 59 EISKEKLNPILODTKKGLREYVNCFPHGIIYNGAFPTMEDPNTHEETAKGND 118  
 59 EIKTKPLNFKODVKKGLRYANLEPKYIIMNGAIPQWEPBGNDKHTGCCGND 118  
 Oy 119 PLVCEIGSKVARGELIGVGLIAMDDEGETMDKLVAINVDDPDANNDINDVYRL 178  
 119 PLVCEIGSKVARGELIGVGLIAMDDEGETMDKLVAINVDDPDANNDINDVYRL 178  
 Db 178 MGLLRATNEMERYIKIPDGKPENOFAFSGECKNKKYAEVIOCREAMEKLIIFKTSBG 237  
 178 KPGYLEATVWFRYKVPDCKPENEFANAEFKDFAIDIIKSTHDMKALVYTKTKNG 238  
 Oy 238 EIDLNTLTGTP-SFSPSATSAPTA-SPA-APA-KIDQSIDKWF 279  
 239 GISCMNTLSESPKCDPDARAIVALLPPCHSACTVPTVDYDKW 284

RESULT 3 PRELIMINARY: PRT: 286 AA.  
 ID 087118;  
 AC 087118;  
 DT 01-JUL-1997 (TREMBLREL. 04, CREATED)  
 DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)  
 DT 01-JUL-1997 (TREMBLREL. 04, LAST ANNOTATION UPDATE)  
 DE HYPOHETICAL 32.9 KD PROTEIN.  
 GN SPAC3A12.02.  
 OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).  
 CC EURARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCYCETES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972H-;  
 RA BADOCK K., CHURCHER C.M.;

RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972H-;  
 RA WOOD V., BARRELL B.G., RAJANDREAM M.A.;  
 RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL; 295395; E316097;  
 KW HYPOHETICAL PROTEIN.  
 SQ SEQUENCE 286 AA; 32898 MW; E702438D CRC32;

Query Match 40.5%; Score 864; DB 1; Length 286;  
 Best Local Similarity 55.5%; Pred. No. 1.72e-164;  
 Matches 122; Conservative 40; Mismatches 55; Indels 3; Gaps 3;  
 Db 23 DFRVYCYKNNK-P-ISEFHDVPLTSKDPENFWTEIPMTQAKCEISLSPFHPKODLK 80  
 16 EYRVF-LKNKGGYISPFPHDIPYADKDVFNHNVVEPRSNKAKMELATKODPLNPKODV 74  
 Db 81 NGKLRYANSPYHGFIMWYGAIPQWEPNVIDSTKMGDPLDVCIGSGYIGY 140  
 75 KGLRYANLEPKYIIMNGAIPQWEPBGNDKHTGCCGNDPDIVCIGSKVARGE 134  
 Db 141 IKOVYIGALGLDQGETMDKLVAINVDDPDANNDINDVYKRLKPGYLEATVWFRYK 200  
 135 IIGVYIGLIMDEGETMDKLVAINVDDPDANNDINDVYKRLKPGYLEATVWFRYK 194  
 Db 201 IPDGKPNRFFEGDNYLPKSDALDIQAOCQHWKVSARDK 240  
 195 VPDGKPNRFFANAEFKDFAIDIIKSTHDMKALVYTK 234

RESULT 4 PRELIMINARY: PRT: 164 AA.  
 ID 005724;  
 AC 005724;  
 DT 01-JUL-1997 (TREMBLREL. 04, CREATED)  
 DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)  
 DT 01-JUL-1997 (TREMBLREL. 04, LAST ANNOTATION UPDATE)  
 DE PYROPHOSPHATASE (FRAGMENT).  
 GN PMK2PPA.  
 OS BACILLUS STEAROTHERMOPHILUS.  
 CC PROCAROTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 12016;  
 RA SATOH T., ISHII K., KOTAMA M., SAKURAI N., KAJI H., HACHIMORI A.,  
 RA IRIE M., SAMEJIMA T.;  
 RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL; AB003087; D1020624;  
 FT NON\_TER 1  
 NON\_TER 164  
 SQ SEQUENCE 164 AA; 18796 MW; BD71D735 CRC32;

Query Match 7.2%; Score 154; DB 9; Length 164;  
 Best Local Similarity 33.3%; Pred. No. 8.78e-10;  
 Matches 28; Conservative 18; Mismatches 35; Indels 3; Gaps 1;

Db 56 DGGPDIVTTNTPPGVIVDERVGYIMVDSGEEDAKLIGVYEDP---REDEVERSI 112  
 116 DNDPDIVCIGSKVARGELIGVGLIAMDDEGETMDKLVAINVDDPDANNDINDV 175  
 Db 113 EDLPQKLEIAFFERYKDLCK 136  
 176 KRLKPGYLEATVWFRYKVPDCK 199

RESULT 5 PRELIMINARY: PRT: 215 AA.  
 ID 023979;  
 AC 023979;  
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)  
 DE SOLUBLE INORGANIC PYROPHOSPHATASE (EC 3.6.1.1).  
 GN IIP.



[illegible]

RESULT	8	PRELIMINARY;	PT;	214	AA.
ID	O22537				
AC	O22537;				
DT	01-JAN-1998 (TREMBL:REL_05, CREATED)				
DT	01-JAN-1998 (TREMBL:REL_05, LAST SEQUENCE UPDATE)				
DT	01-JAN-1998 (TREMBL:REL_05, LAST ANNOTATION UPDATE)				
DE	INORGANIC PYROPHOSPHATASE.				
GN	IPP.				
OS	ORYZA SATIVA (RICE).				
OC	EURAROTIA; PLANTAE; EMBRYOPHYTA; ANGIOSPERMAE; MONOCOTYLEDONEAE;				
CC	CYPERALES; GRAMINEAE.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-MILYANG 23; TISSUE-SEED;				
RA	LEE M.C.; KIM C.S.; EUN M.Y.;				
RL	SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.				
DR	EMBL, AF022733; G2370501; --				
SO	SEQUENCE	214	AA;	23610	MM; ED159571
					CRC32;

QY	116	DNDP	IVCEIGSKVCAREBIIIGVKVLGILAMIDEGETDMKVI	AINVDDPDAA	NYNDINDV	175
DB	126	-P-EHN-LKEIANFET	YKKLEGG	146		
QY	176	KRLKPGYLEATVDMFR	RYKVPDGG	199		
RESULT	7					
ID	049071					
AC	Q49071	PRELIMINARY	PRT	136	AA	

SUBMITTED (NOV-1994) TO EMBL/GENBANK/DDJ DATA BANKS.  
 CC -1- SUBUNIT: HEXAMER, EACH SUBUNIT IS COMPOSED OF AN ACIDIC AND A  
 CC BASIC CHAIN DERIVED FROM A SINGLE PRECURSOR AND LINKED BY A  
 CC DISULFIDE BOND (BY SIMILARITY).  
 CC -1- SIMILARITY: TO OTHER ITS SEED STORAGE PROTEINS (GLOBULINS).  
 CC EMBL: Z46803; G600108;  
 DR PROSITE; PS00305; 118; SEED\_STORAGE; 1.  
 DR SIGNAL.  
 FT SIGNAL.  
 FT SEQUENCE 1 21 POTENTIAL.  
 SQ SEQUENCE 482 AA; 54667 MW; 901C5103 CRC32;  
 Query Match 5.18; Score 109; DB 8; Length 482;  
 Best Local Similarity 32.33; Pred. No. 1,788-02;  
 Matches 30; Conservative 21; Mismatches 33; Indels 9; Gaps 8;  
 Db 138 GDIIVPGNVLMYNNQDPT- VIALSLDTGSSN-NQDQIR-R-FYLAGNOGEFFL 192  
 133 GELIVKVLGLAMIDGETDMKVIINVDPPAANYNDINDVKRLKPGYLEAIVDM-FR 191  
 193 FYOREGGKQKQENDGNNTSGFROPLEALN 225  
 192 RY-KVDDGKPENEFAFNAEFD-K-DPAIDIIK 221  
 RESULT 12 PRELIMINARY; PR; 356 AA.  
 ID Q30444;  
 AC Q30444;  
 DT 01-NOV-1996 (TREMBLREL. 01. CREATED)  
 DT 01-NOV-1996 (TREMBLREL. 01. LAST SEQUENCE UPDATE)  
 DT 01-NOV-1996 (TREMBLREL. 01. LAST ANNOTATION UPDATE)  
 DT MHC CLASS I CAVA-G\*04 (FRAGMENT).  
 CN CAVA-G.  
 GN CALITHRIX JACCHUS (COMMON MARMOSET).  
 OS EURACOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA.  
 OC EUHIERIA; PRIMATES.  
 RN ([1]  
 RP SEQUENCE FROM N.A.  
 RA CADAVID L.F., SHUFFLEBOTHAM C., RUIZ F.J., YEAGER M., HUGHES A.L.,  
 RA WATKINS D.I.;  
 RL SUBMITTED (JUN-1996) TO EMBL/GENBANK/DDJ DATA BANKS.  
 DR EMBL; U596440; G1389927;  
 KM MHC.  
 FT NON TER 1  
 SQ SEQUENCE 356 AA; 39894 MW; ECFE4FAE CRC32;  
 Query Match 5.08; Score 106; DB 5; Length 356;  
 Best Local Similarity 27.88; Pred. No. 4,878-02;  
 Matches 15; Conservative 15; Mismatches 20; Indels 4; Gaps 4;  
 Db 142 LINEDLSWTAIPVAAQIQR-KMAANAEAR-RAYLEGTCEWELRRY-LENGK 192  
 147 IDGETDMKVIINVDPPAANYNDINDVKRLKPGYLEA-VDMFRRYKVPDGK 199  
 RESULT 13 PRELIMINARY; PR; 357 AA.  
 ID Q30914;  
 AC Q30914;  
 DT 01-NOV-1996 (TREMBLREL. 01. CREATED)  
 DT 01-NOV-1996 (TREMBLREL. 01. LAST SEQUENCE UPDATE)  
 DT 01-JAN-1998 (TREMBLREL. 05. LAST ANNOTATION UPDATE)  
 DT MHC CLASS I A (FRAGMENT).  
 DE PAN TROGLODYTES (CHIMPANZEE).  
 OS EURACOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA.  
 OC EUHIERIA; PRIMATES.  
 RN ([1]  
 RP SEQUENCE FROM N.A.  
 RA MCADAM S.N., BOSON J.E., LIU X., GARBER T.L., HUGHES A.L.,  
 RA BONTROP R.E., WATKINS D.I.;  
 RL J. IMMUNOL. 154:6421-6429(1995).  
 DR EMBL; U10539; G506670;  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KM MHC.

```

FT  NON TER      1      1
   SEQUENCE      357 AA; 39890 MM; 24EFILF7 CRC32;

Query Match
Best Local Similarity 27.8%; Score 107; DB 5; Length 357;
Matches 15; Conservative 15; Mismatches 20; Indels 4; Gaps 4;

DB  142 LNEIDLRSTADMAAQITQR-KWEAAHAERLRA-YLCTGVEMLRVY-LENGK 192
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  147 IDEGETDMKVIAINVDPPDPAANYNDINDVKRLKPGYLEAT-VDMEFRYKVPDOK 199

RESULT  14
ID  Q08680; PRELIMINARY; PRT; 367 AA.
AC  Q08680;
DT  01-NOV-1996 (TREMBLERL. 01, CREATED)
DT  01-NOV-1996 (TREMBLERL. 01, LAST SEQUENCE UPDATE)
DE  01-JAN-1998 (TREMBLERL. 05, LAST ANNOTATION UPDATE)
DE  RT1 CLASS I HISTOCOMPATIBILITY ANTIGEN, LM2 ALPHA CHAIN PRECURSOR.
GN  LM2.
OS  RATIUS NORVEGICUS (RAT).
OC  EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC  EUTHERIA; RODENTIA.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-LEW/GUN; TISSUE-CONA STIMULATED LYMPHOCYTES;
RX  MEDLINE: 94222444.
RA  WALTER L., HEINE L., GUENTHER E.;
RL  IMMUNOGENETICS 39:351-354(1994).
DR  EMBL: X70066; G56610;
KW  MHC I; TRANSMEMBRANE.
FT  SIGNAL 1 21 POTENTIAL.
FT  CHAIN 22 367 RT1 CLASS I HISTOCOMPATIBILITY ANTIGEN,
SQ  SEQUENCE 367 AA; 41053 MM; 60475057 CRC32; LM2 ALPHA CHAIN.

Query Match
Best Local Similarity 27.0%; Score 107; DB 5; Length 367;
Matches 20; Conservative 19; Mismatches 29; Indels 6; Gaps 6;

DB  122 CDVGSGLRIGYDQFAIDGRYIALNEDLKT-WMA-ADTAQMTRNKWDQGEARHKA 179
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  123 CEIGSKVCA-RG-EIIGKVILGILAMIDEGTDMKVIAINVDPPDPAANYNDINDVKRLK 180
DB  180 -YLOGTCVEMLRVY 192
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  181 GYLEAT-VDMEFRY 193

RESULT  15
ID  019445; PRELIMINARY; PRT; 377 AA.
AC  019445;
DT  01-JAN-1998 (TREMBLERL. 05, CREATED)
DT  01-JAN-1998 (TREMBLERL. 05, LAST SEQUENCE UPDATE)
DE  01-JAN-1998 (TREMBLERL. 05, LAST ANNOTATION UPDATE)
DE  MHC CLASS IB ANTIGEN.
GN  RT1.CI.
OS  RATIUS NORVEGICUS (RAT).
OC  EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC  EUTHERIA; RODENTIA.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-LEW;
RX  MEDLINE: 95278971.
RA  LAMBRACHT D., WONIGEIT K.;
RL  IMMUNOGENETICS 41:375-379(1995).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN-LEW;
RA  LAMBRACHT D., DUEVEL H., WONIGEIT K.;
RL  SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN  [3]
RP  SEQUENCE FROM N.A.

```

```

RC  STRAIN-LEW;
RA  LAMBRACHT D.;
RL  SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
SQ  SEQUENCE 377 AA; 42200 MM; 335C92C1 CRC32;

Query Match
Best Local Similarity 27.0%; Score 107; DB 5; Length 377;
Matches 20; Conservative 19; Mismatches 29; Indels 6; Gaps 6;

DB  122 CDVGSGLRIGYDQFAIDGRYIALNEDLKT-WMA-ADTAQMTRNKWDQGEARHKA 179
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  123 CEIGSKVCA-RG-EIIGKVILGILAMIDEGTDMKVIAINVDPPDPAANYNDINDVKRLK 180
DB  180 -YLOGTCVEMLRVY 192
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  181 GYLEAT-VDMEFRY 193

```

Search completed: Thu Sep 17 16:57:04 1998  
Job time : 124 secs.

